

Mouse Dnmt3a DNA sequence

1 GAATTCCGGC CTGCTGCCGG GCCGCCCCGAC CCGCCGGGCC ACACGGCAGA  
51 GCCGCCTGAA GCCCAGCGCT GAGGCTGCAC TTTTCCGAGG GCTTGACATC  
101 AGGGTCTATG TTTAAGTCTT AGCTCTTGCT TACAAAGACC ACGGCAATTC  
151 CTTCTCTGAA GCCCTCGCAG CCCACACAGG CCCTCGCAGC CCCAGCCTGC  
201 CGCCTACTGC CCAGCAATGC CCTCCAGCGG CCCCGGGGAC ACCAGCAGCT  
251 CCTCTCTGGA GCGGGAGGAT GATCGAAAGG AAGGAGAGGA ACAGGAGGAG  
301 AACCGTGGA AGGAAGAGCG CCAGGAGCCC AGCGCCACGG CCCGGAAGGT  
351 GGGGAGGCCT GGCCGAAGC GCAAGCAGCC ACCGGTGGA AGCAGTGACA  
401 CCCCCAAGGA CCCAGCAGTG ACCACCAAGT CTCAGCCCAT GGGCCAGGAC  
451 TCTGGCCCCT CAGATCTGCT ACCCAATGGA GACTTGGAGA AGCGGAGTGA  
501 ACCCCAACCT GAGGAGGGGA GCCCAGCTGC AGGGCAGAAG GGTGGGGCCC  
551 CAGCTGAAGG AGAGGGAAGT GAGACCCAC CAGAAGCCTC CAGAGCTGTG  
601 GAGAATGGCT GCTGTGTGAC CAAGGAAGGC CGTGGAGCCT CTGCAGGAGA  
651 GGGCAAAGAA CAGAAGCAGA CCAACATCGA ATCCATGAAA ATGGAGGGCT  
701 CCCGGGGCCG ACTGCGAGGT GGCTTGGCT GGGAGTCCAG CCTCCGTCAG  
751 CGACCCATGC CAAGACTCAC CTTCCAGGCA GGGGACCCCT ACTACATCAG  
801 CAAACGAAA CGGGATGAGT GGCTGCCACG TTGAAAAGG GAGGCTGAGA  
851 AGAAAGCCAA GGTAATTGCA GTAATGAATG CTGTGGAAGA GAACCAGGCC  
901 TCTGGAGAGT CTCAGAAGGT GGAGGAGGCC AGCCCTCCTG CTGTGCAGCA  
951 GCCCACGGAC CCTGCTTCTC CGACTGTGGC CACCACCCCT GAGCCAGTAG  
1001 GAGGGGATGC TGGGACAAG AATGCTACCA AAGCAGCCGA CGATGAGCCT  
1051 GAGTATGAGG ATGGCCGGGG CTTTGGCATT GGAGAGCTGG TGTGGGGGAA  
1101 ACTTCGGGGC TTCTCCTGGT GGCCAGGCCG AATTGTGTCT TGGTGGATGA

FIG. 1A-1

1151 CAGGCCGGAG CCGAGCAGCT GAAGGCACTC GCTGGGTCAT GTGGTTCGGA  
 1201 GATGGCAAGT TCTCAGTGGT GTGTGTGGAG AAGCTCATGC CGCTGAGCTC  
 1251 CTTCTGCAGT GCATTCCACC AGGCCACCTA CAACAAGCAG CCCATGTACC  
 1301 GCAAAGCCAT CTACGAAGTC CTCCAGGTGG CCAGCAGCCG TGCCGGGAAG  
 1351 CTGTTTCCAG CTTGCCATGA CAGTGATGAA AGTGACAGTG GCAAGGCTGT  
 1401 GGAAGTGCAG AACAAGCAGA TGATTGAATG GGCCCTCGGT GGCTTCCAGC  
 1451 CCTCGGGTCC TAAGGGCCTG GAGCCACCAG AAGAAGAGAA GAATCCTTAC  
 1501 AAGGAAGTTT ACACCGACAT GTGGGTGGAG CCTGAAGCAG CTGCTTACGC  
 1551 CCCACCCCCA CCAGCCAAGA AACCCAGAAA GAGCACAACA GAGAAACCTA  
 1601 AGGTCAAGGA GATCATTGAT GAGCGCACAA GGGAGCGGCT GGTGTATGAG  
 1651 GTGCGCCAGA AGTGCAGAAA CATCGAGGAC ATTTGTATCT CATGTGGGAG  
 1701 CCTCAATGTC ACCCTGGAGC ACCCACTCTT CATTGGAGGC ATGTGCCAGA  
 1751 ACTGTAAGAA CTGCTTCTTG GAGTGTGCTT ACCAGTATGA CGACGATGGG  
 1801 TACCAGTCCT ATTGCACCAT CTGCTGTGGG GGGCGTGAAG TGCTCATGTG  
 1851 TGGGAACAAC AACTGCTGCA GGTGCTTTTG TGTCGAGTGT GTGGATCTCT  
 1901 TGGTGGGGCC AGGAGCTGCT CAGGCAGCCA TTAAGGAAGA CCCCTGGAAC  
 1951 TGCTACATGT GCGGGCATAA GGGCACCTAT GGGCTGCTGC GAAGACGGGA  
 2001 AGACTGGCCT TCTCGACTCC AGATGTTCTT TGCCAATAAC CATGACCAGG  
 2051 AATTTGACCC CCCAAAGGTT TACCCACCTG TGCCAGCTGA GAAGAGGAAG  
 2101 CCCATCCGCG TGCTGTCTCT CTTTGATGGG ATTGCTACAG GGCTCCTGGT  
 2151 GCTGAAGGAC CTGGGCATCC AAGTGGACCG CTACATTGCC TCCGAGGTGT  
 2201 GTGAGGACTC CATCACGGTG GGCATGGTGC GGCACCAGGG AAAGATCATG  
 2251 TACGTGGGGG ACGTCCGCAG CGTCACACAG AAGCATATCC AGGAGTGGGG  
 2301 CCCATTGCAC CTGGTGATTG GAGGCAGTCC CTGCAATGAC CTCTCCATTG

FIG. 1A-2

2351 TCAACCCTGC CCGCAAGGGA CTTTATGAGG GTACTGGCCG CCTCTTCTTT  
 2401 GAGTTCTACC GCCTCCTGCA TGATGCGCGG CCCAAGGAGG GAGATGATCG  
 2451 CCCCTTCTTC TGGCTCTTTG AGAATGTGGT GGCCATGGGC GTTAGTGACA  
 2501 AGAGGGACAT CTCGCGATTT CTTGAGTCTA ACCCCGTGAT GATTGACGCC  
 2551 AAAGAAGTGT CTGCTGCACA CAGGGCCCGT TACTTCTGGG GTAACCTTCC  
 2601 TGGCATGAAC AGGCCTTTGG CATCCACTGT GAATGATAAG CTGGAGCTGC  
 2651 AAGAGTGTCT GGAGCACGGC AGAATAGCCA AGTTCAGCAA AGTGAGGACC  
 2701 ATTACCACCA GGTCAAATC TATAAAGCAG GGCAAAGACC AGCATTTCCT  
 2751 CGTCTTCATG AACGAGAAGG AGGACATCCT GTGGTGCACT GAAATGGAAA  
 2801 GGGTGTITGG CTTCCCCGTC CACTACACAG ACGTCTCCAA CATGAGCCGC  
 2851 TTGGCGAGGC AGAGACTGCT GGGCCGATCG TGGAGCGTGC CGGTCATCCG  
 2901 CCACCTCTTC GCTCCGCTGA AGGAATATTT TGCTTGTTG TAAGGGACAT  
 2951 GGGGGCAAAC TGAAGTAGTG ATGATAAAAA AGTTAAACAA ACAAAACAAAC  
 3001 AAAAAACAAA ACAAACAAT AAAACACCAA GAACGAGAGG ACGGAGAAAA  
 3051 GTTCAGCACC CAGAAGAGAA AAAGGAATTT AAAGCAAACC ACAGAGGAGG  
 3101 AAAACGCCCG AGGGCTTGGC CTTGCAAAAG GGTGGACAT CATCTCCTGA  
 3151 GTTTTCAATG TTAACCTTCA GTCCTATCTA AAAAGCAAAA TAGGCCCTC  
 3201 CCCTTCTTCC CCTCCGGTCC TAGGAGGCGA ACTTTTGTG TTCTACTCTT  
 3251 TTTCAGAGGG GTTTTCTGTT TGTTTGGTT TTTGTTTCTT GCTGTGACTG  
 3301 AAACAAGAGA GTTATTGCAG CAAAATCAGT AACAAACAAA AGTAGAAATG  
 3351 CCTTGGAGAG GAAAGGGAGA GAGGGAAAAT TCTATAAAAA CTTAAATAT  
 3401 TGGTTTTTTT TTTTTTCCT TTTCTATATA TCTCTTGGT TGTCTCTAGC  
 3451 CTGATCAGAT AGGAGCACAA ACAGGAAGAG AATAGAGACC CTCGGAGGCA  
 3501 GAGTCTCTC TCCCACCCCC CGAGCAGTCT CAACAGCACC ATTCCTGGTC

FIG. 1A-3

3551 ATGCAAAACA GAACCCAACT AGCAGCAGGG CGCTGAGAGA ACACCACACC  
3601 AGACACTTTC TACAGTATTT CAGGTGCCTA CCACACAGGA AACCTTGAAG  
3651 AAAACCAGTT TCTAGAAGCC GCTGTTACCT CTTGTTTACA GTTTATATAT  
3701 ATATGATAGA TATGAGATAT ATATATATAA AAGGTACTGT TAACTACTGT  
3751 ACATCCCGAC TTCATAATGG TGCTTTCAAA ACAGCGAGAT GAGCAAAGAC  
3801 ATCAGCTTCC GCCTGGCCCT CTGTGCAAAG GGTTTCAGCC CAGGATGGGG  
3851 AGAGGGGAGC AGCTGGAGGG GGTTTAAACA AACTGAAGGA TGACCCATAT  
3901 CACCCCCCAC CCCTGCCCCA TGCCTAGCTT CACCTGCCAA AAAGGGGCTC  
3951 AGCTGAGGTG GTCGGACCCT GGGGAAGCTG AGTGTGGAAT TTATCCAGAC  
4001 TCGCGTGCAA TAACCTTAGA ATATGAATCT AAAATGACTG CCTCAGAAAA  
4051 ATGGCTTGAG AAAACATTGT CCCTGATTTT GAATTGTC A GCCACGTTGA  
4101 AGGCCCCCTG TGGGATCAGA AATATTCCAG AGTGAGGGAA AGTGACCCGC  
4151 CATTAACCCC NCCTGGAGCA AATAAAAAAA CATACAAAAT GT

FIG. 1A-4

Mouse Dnmt3b1 DNA Sequence

1 GAATTCGGG CGCCGGGGTT AAGCGGCCA AGTAAACGTA GCGCAGCGAT  
51 CGGCGCCGGA GATTGCGAA CCCGACACTC CGGCGCGCCC GCCGGCCAGG  
101 ACCCGCGGCG CGATCGCGGC GCCGCGCTAC AGCCAGCCTC ACGACAGGCC  
151 CGCTGAGGCT TGTGCCAGAC CTTGGAAACC TCAGGTATAT ACCTTTCCAG  
201 ACGCGGGATC TCCCCTCCCC CATCCATAGT GCCTTGGGAC CAAATCCAGG  
251 GCCTTCTTTC AGGAAACAAT GAAGGGAGAC AGCAGACATC TGAATGAAGA  
301 AGAGGGTGCC AGCGGGTATG AGGAGTGCAT TATCGTTAAT GGGAACTTCA  
351 GTGACCAGTC CTCAGACACG AAGGATGCTC CCTCACCCCC AGTCTTGGAG  
401 GCAATCTGCA CAGAGCCAGT CTGCACACCA GAGACCAGAG GCCGCAGGTC  
451 AAGCTCCCGG CTGTCTAAGA GGGAGGTCTC CAGCCTTCTG AATTACACGC  
501 AGGACATGAC AGGAGATGGA GACAGAGATG ATGAAGTAGA TGATGGGAAT  
551 GGCTCTGATA TTCTAATGCC AAAGCTCACC CGTGAGACCA AGGACACCAG  
601 GACGCGCTCT GAAAGCCCGG CTGTCCGAAC CCGACATAGC AATGGGACCT  
651 CCAGCTTGGA GAGGCAAAGA GCCTCCCCCA GAATCACCCG AGGTCCGCAG  
701 GGCCGCCACC ATGTGCAGGA GTACCCTGTG GAGTTTCCGG CTACCAGGTC  
751 TCGGAGACGT CGAGCATCGT CTTAGCAAG CACGCCATGG TCATCCCCTG  
801 CCAGCGTCGA CTTATGGAA GAAGTGACAC CTAAGAGCGT CAGTACCCCA  
851 TCAGTTGACT TGAGCCAGGA TGGAGATCAG GAGGGTATGG ATACCACACA  
901 GGTGGATGCA GAGAGCAGAG ATGGAGACAG CACAGAGTAT CAGGATGATA  
951 AAGAGTTTGG AATAGGTGAC CTCGTGTGGG GAAAGATCAA GGGCTTCTCC  
1001 TGGTGGCCTG CCATGGTGGT GTCCTGAAA GCCACCTCCA AGCGACAGGC

FIG. 1B-1

1051 CATCCCCGA ATGCGCTGGG TACAGTGGTT TGGTGATGGC AAGTTTTCTG  
 1101 AGATCTCTGC TGACAACTG GTGGCTCTGG GGCTGTTGAG CCAGCACTTT  
 1151 AATCTGGCTA CCTTCAATAA GCTGGTTTCT TATAGGAAGG CCATGTACCA  
 1201 CACTCTGGAG AAAGCCAGGG TTCGAGCTGG CAAGACCTTC TCCAGCAGTC  
 1251 CTGGAGAGTC ACTGGAGGAC CAGCTGAAGC CCATGCTGGA GTGGGCCCAC  
 1301 GGTGGCTTCA AGCCTACTGG GATCGAGGGC CTCAAACCCA ACAAGAAGCA  
 1351 ACCAGTGGTT AATAAGTCGA AGGTGCGTCG TTCAGACAGT AGGAACTTAG  
 1401 AATCCAGGAG ACGCGAGAAC AAAAGTCGAA GACGCACAAC CAATGACTCT  
 1451 GCTGCTTCTG AGTCCCCCCC ACCCAAGCGC CTCAAGACAA ATAGCTATGG  
 1501 CGGAAGGAC CGAGGGGAGG ATGAGGAGAG CCGAGAACGG ATGGCTTCTG  
 1551 AAGTACCAA CAACAAGGGC AATCTGGAAG ACCGCTGTTT GTCCTGTGGA  
 1601 AAGAAGAACC CTGTGTCCTT CCACCCCTC TTTGAGGGTG GGCTCTGTCA  
 1651 GAGTTGCCGG GATCGCTTCC TAGAGCTCTT CTACATGTAT GATGAGGACG  
 1701 GCTATCAGTC CTA CTGCTGACC GTGTGCTGTG AGGGCCGTGA ACTGCTGCTG  
 1751 TGCAGTAACA CAAGCTGCTG CAGATGCTTC TGTGTGGAGT GTCTGGAGGT  
 1801 GCTGGTGGGC GCAGGCACAG CTGAGGATGC CAAGCTGCAG GAACCCTGGA  
 1851 GCTGCTATAT GTGCCTCCCT CAGCGCTGCC ATGGGGTCCT CCGACGCAGG  
 1901 AAAGATTGGA ACATGCGCCT GCAAGACTTC TTCACTACTG ATCCTGACCT  
 1951 GGAAGAATTT GAGCCACCCA AGTTGTACCC AGCAATTCCT GCAGCCAAAA  
 2001 GGAGGCCCAT TAGAGTCCTG TCTCTGTTTG ATGGAATTGC AACGGGTAC  
 2051 TTGGTGCTCA AGGAGTTGGG TATTAAAGTG GAAAAGTACA TTGCCTCCGA  
 2101 AGTCTGTGCA GAGTCCATCG CTGTGGGAAC TGTTAAGCAT GAAGGCCAGA  
 2151 TCAAATATGT CAATGACGTC CGGAAAATCA CCAAGAAAAA TATTGAAGAG  
 2201 TGGGGCCCGT TCGACTTGGT GATTGGTGGA AGCCCATGCA ATGATCTCTC

FIG. 1B-2

2251 TAACGTCAAT CCTGCCCCGA AAGGTTTATA TGAGGGCACA GGAAGGCTCT  
 2301 TCTTCGAGTT TTACCACTTG CTGAATTATA CCCGCCCCAA GGAGGGCGAC  
 2351 AACCGTCCAT TCTTCTGGAT GTTCGAGAAT GTTGTGGCCA TGAAAGTGAA  
 2401 TGACAAGAAA GACATCTCAA GATTCTGGC ATGTAACCCA GTGATGATCG  
 2451 ATGCCATCAA GGTGTCTGCT GCTCACAGGG CCCGGTACTT CTGGGGTAAC  
 2501 CTACCCGGAA TGAACAGGCC CGTGATGGCT TCAAAGAATG ATAAGCTCGA  
 2551 GCTGCAGGAC TGCCTGGAGT TCAGTAGGAC AGCAAAGTTA AAGAAAGTGC  
 2601 AGACAATAAC CACCAAGTCG AACTCCATCA GACAGGGCAA AAACCAGCTT  
 2651 TTCCCTGTAG TCATGAATGG CAAGGACGAC GTTTTGTGGT GCACTGAGCT  
 2701 CGAAAGGATC TTCGGCTTCC CTGCTCACTA CACGGACGTG TCCAACATGG  
 2751 GCCGCGGGCG CCGTCAGAAG CTGCTGGGCA GGTCTGGAG TGTACCGGTC  
 2801 ATCAGACACC TGTTCGCCCC CTTGAAGGAC TACTTTGCCT GTGAATAGTT  
 2851 CTACCCAGGA CTGGGGAGCT CTCGGTCAGA GCCAGTGCCC AGAGTCACCC  
 2901 CTCCCTGAAG GCACCTCACC TGTCCCCTTT TTAGCTCACC TGTGTGGGGC  
 2951 CTCACATCAC TGTACCTCAG CTTTCTCCTG CTCAGTGGGA GCAGAGCCTC  
 3001 CTGGCCCTTG CAGGGGAGCC CCGGTGCTCC CTCCGTGTGC ACAGCTCAGA  
 3051 CCTGGCTGCT TAGAGTAGCC CGGCATGGTG CTCATGTTCT CTTACCCTGA  
 3101 AACTTTAAAA CTTGAAGTAG GTAGTAAGAT GGCTTTCTTT TACCCTCCTG  
 3151 AGTTTATCAC TCAGAAGTGA TGGCTAAGAT ACCAAAAAAA CAAACAAAAA  
 3201 CAGAAACAAA AAACAAAAAA AAACCTCAAC AGCTCTCTTA GTACTCAGGT  
 3251 TCATGCTGCA AAATCACTTG AGATTTTGTT TTTAAGTAAC CCGTGCTCCA  
 3301 CATTTGCTGG AGGATGCTAT TGTGAATGTG GGCTCAGATG AGCAAGGTCA  
 3351 AGGGGCCAAA AAAAATTCCC CCTCTCCCCC CAGGAGTATT TGAAGATGAT  
 3401 GTTTATGGTT TAAGTCTTCC TGGCACCTTC CCCTTGCTTT GGTACAAGGG

FIG. 1B-3

3451 CTGAAGTCCT GTTGGTCTTG TAGCATTTC CAGGATGATG ATGTCAGCAG  
3501 GGATGACATC ACCACCTTTA GGGCTTTTCC CTGGCAGGGG CCCATGTGGC  
3551 TAGTCCTCAC GAAGACTGGA GTAGAATGTT TGGAGCTCAG GAAGGGTGGG  
3601 TGGAGTGGCC CTCTTCAGG TGTGAGGGAT ACGAAGGAGG AAGCTTAGGG  
3651 AAATCCATTG CCCACTCCCT CTGCCAAAT GAGGGGCCCA GTCCCCAACA  
3701 GCTCAGGTCC CCAGAACCCC CTAGTTCCTC ATGAGAAGCT AGGACCAGAA  
3751 GCACATCGTT CCCCTTATCT GAGCAGTGTT TGGGGAAC TAAGTAAAAC  
3801 CTTCTGGAGA TGTTAAAAGC TTTTACCCC ACGATAGATT GTGTTTTTAA  
3851 GGGGTGCTTT TTTAGGGGC ATCACTGGAG ATAAGAAAGC TGCATTTTCA  
3901 AAATGCCATC GTAATGGTTT TTAACACCT TTTACCTAAT TACAGGTGCT  
3951 ATTTTATAGA AGCAGACAAC ACTTCTTTTT ATGACTCTCA GACTTCTATT  
4001 TTCATGTTAC CATTTTTTTT GTAACGCA AGGTGTGGGC TTTGTAACT  
4051 TCACAGGTGT GGGGAGAGAC TGCCTTGTTT CAACAGTTTG TCTCCACTGG  
4101 TTTCTAATTT TTAGGTGCAA AGATGACAGA TGCCAGAGT TTACCTTTCT  
4151 GGTGATTAA AGTTGTATTT CTCTAAAAA AAAAAAAAAA AAAAA

FIG. 1B-4

# Human DNMT3A DNA Sequence

1 GCGCGG CACCAGGGCG CGCAGCCGGG  
28 CCGGCCCGAC CCCACCGGCC ATACGGTGA GCCATCGAAG CCCCCACCCA  
78 CAGGCTGACA GAGGCACCGT TCACCAGAGG GCTCAACACC GGGATCTATG  
128 TTTAAGTTTT AACTCTCGCC TCCAAAGACC ACGATAATTC CTTCCCCAAA  
178 GCCCAGCAGC CCCCCAGCCC CGCGCAGCCC CAGCCTGCCT CCCGGCGCCC  
228 AGATGCCCCG CATGCCCTCC AGCGGCCCCG GGGACACCAG CAGCTCTGCT  
278 GCGGAGCGGG AGGAGGACCG AAAGGACGGA GAGGAGCAGG AGGAGCCGGG  
328 TGGCAAGGAG GAGCGCCAAG AGCCCAGCAC CACGGCACGG AAGGTGGGGC  
378 GGCCTGGGAG GAAGCGCAAG CACCCCCCGG TGGAAAGCGG TGACACGCCA  
428 AAGGACCCTG CGGTGATCTC CAAGTCCCCA TCCATGGCCC AGGACTCAGG  
478 CGCCTCAGAG CTATTACCCA ATGGGGACTT GGAGAAGCGG AGTGAGCCCC  
528 AGCCAGAGGA GGGGAGCCCT GCTGGGGGGC AGAAGGGCGG GGCCCCAGCA  
578 GAGGGAGAGG GTGCAGCTGA GACCCTGCCT GAAGCCTCAA GAGCAGTGA  
628 AAATGGCTGC TGCACCCCCA AGGAGGGCCG AGGAGCCCCT GCAGAAGCGG  
678 GCAAAGAACA GAAGGAGACC AACATCGAAT CCATGAAAAT GGAGGGCTCC  
728 CGGGGCCGGC TCGGGGTGG CTTGGGCTGG GAGTCCAGCC TCCGTCAGCG  
778 GCCCATGCCG AGGCTCACCT TCCAGGCGGG GGACCCCTAC TACATCAGCA  
828 AGCGCAAGCG GCACGAGTGG CTGGCAGCT GGAAGGGA GGCTGAGAAG  
878 AAAGCCAAGG TCAGTGCAGG AATGAATGCT GTGAAGAAA ACCAGGGGCC  
928 CGGGGAGTCT CAGAAGGTGG AGGAGGCCAG CCCTCCTGCT GTGCAGCAGC  
978 CCACTGACCC CGCATCCCC ACTGTGGCTA CCACGCCTGA GCCCGTGGGG  
1028 TCCGATGCTG GGGACAAGAA TGCCACCAA GCAGGCGATG ACGAGCCAGA

FIG. 1C-1

1078 GTACGAGGAC GGCCGGGGCT TTGGCATTGG GGAGCTGGTG TGGGGGAAAC  
1128 TGCGGGGCTT CTCCTGGTGG CCAGGCCGCA TTGTGTCTTG GTGGATGACG  
1178 GGCCGGAGCC GAGCAGCTGA AGGCACCCGC TGGGTCATGT GGTTCGGAGA  
1228 CGGCAAATTC TCAGTGGTGT GTGTTGAGAA GCTGATGCCG CTGAGCTCGT  
1278 TTTGCAGTGC GTTCCACCAG GCCACGTACA ACAAGCAGCC CATGTACCGC  
1328 AAAGCCATCT ACGAGGTCTT GCAGGTGGCC AGCAGCCGCG CGGGGAAGCT  
1378 GTTCCCGGTG TGCCACGACA GCGATGAGAG TGACACTGCC AAGCCGTCG  
1428 AGGTGCAGAA CAAGCCCATG ATTGAATGGG CCCTGGGGGG CTTCCAGCCT  
1478 TCTGGCCCTA AGGCCTGGA GCCACCAGAA GAAGAGAAGA ATCCCTACAA  
1528 AGAAGTGTA ACAGCATGT GGGTGAACC TGAGGCAGCT GCCTACGCAC  
1578 CACCTCCACC AGCCAAAAAG CCCCAGGAAGA GCACAGCGGA GAAGCCCAAG  
1628 GTCAAGGAGA TTATTGATGA GCGACAAGA GAGCGGCTGG TGTACGAGGT  
1678 GCGGCAGAAG TGCCGGAACA TTGAGGACAT CTGCATCTCC TGTGGGAGCC  
1728 TCAATGTTAC CCTGGAACAC CCCCTCTTCG TTGGAGGAAT GTGCCAAAAC  
1778 TGCAAGAACT GCTTTCTGGA GTGTGCGTAC CAGTACGACG ACGACGGCTA  
1828 CCAGTCCTAC TGCACCATCT GCTGTGGGGG CCGTGAGGTG CTCATGTCCG  
1878 GAAACAACAA CTGCTGCAGG TGCTTTTGCG TGGAGTGTGT GGACCTCTTG  
1928 GTGGGGCCCG GGGCTGCCA GGCAGCCATT AAGGAAGACC CCTGGAAGTG  
1978 CTACATGTGC GGGACAAGG GTACCTACGG GCTGCTGCGG CGGCAGAGG  
2028 ACTGGCCCTC CCGGCTCCAG ATGTTCTTCG CTAATAACCA CGACCAGGAA  
2078 TTTGACCCTC CAAAGGTTTA CCCACCTGTC CCAGCTGAGA AGAGGAAGCC  
2128 CATCCGGGTG CTGTCTCTCT TTGATGGAAT CGCTACAGGG CTCCTGGTGC  
2178 TGAAGGACTT GGCATTGAG GTGGACCGCT ACATTGCCTC GGAGGTGTGT

FIG. 1C-2

2228 GAGGACTCCA TCACGGTGGG CATGGTGCGG CACCAGGGGA AGATCATGTA  
 2278 CGTCGGGGAC GTCCGCAGCG TCACACAGAA GCATATCCAG GAGTGGGGCC  
 2328 CATTGATCT GGTGATTGGG GGCAGTCCCT GCAATGACCT CTCCATCGTC  
 2378 AACCTGCTC GCAAGGCCT CTACGAGGGC ACTGGCCGGC TCTTCTTTGA  
 2428 GTTCTACCGC CTCCTGCATG ATGCGCGGCC CAAGGAGGGA GATGATCGCC  
 2478 CCTTCTTCTG GCTCTTTGAG AATGTGGTGG CCATGGGCGT TAGTGACAAG  
 2528 AGGGACATCT CGCGATTCT CGAGTCCAAC CCTGTGATGA TTGATGCCAA  
 2578 AGAAGTGTCA GCTGCACACA GGGCCCGCTA CTTCTGGGT AACCTTCCCG  
 2628 GTATGAACAG GCCGTTGGCA TCCACTGTGA ATGATAAGCT GGAGCTGCAG  
 2678 GAGTGTCTGG AGCATGGCAG GATAGCCAAG TTCAGCAAAG TGAGGACCAT  
 2728 TACTACGAGG TCAAACCTCA TAAAGCAGG CAAAGACCAG CATTTTCCTG  
 2778 TCTTCATGAA TGAGAAAGAG GACATCTTAT GGTGCACTGA AATGGAAAGG  
 2828 GTATTTGGTT TCCCAGTCCA CTATACTGAC GTCTCCAACA TGAGCCGCTT  
 2878 GCGGAGGCAG AGACTGCTGG GCCGGTCATG GAGCGTGCCA GTCATCCGCC  
 2928 ACCTCTTCGC TCCGCTGAAG GAGTATTTTG CGTGTGTGTA AGGGACATGG  
 2978 GGGCAAACCTG AGGTAGCGAC ACAAAGTTAA ACAAACAAAC AAAAAACACA  
 3028 AAACATAATA AAACACCAAG AACATGAGGA TGGAGAGAAG TATCAGCACC  
 3078 CAGAAGAGAA AAAGGAATTT AAAACAAAA CCACAGAGGC GGAAATACCG  
 3128 GAGGGCTTTG CCTTCCGAAA AGGGTTGGAC ATCATCTCCT GATTTTTCAA  
 3178 TGTATTCTT CAGTCCTATT TAAAAACAAA ACCAAGCTCC CTTCCTTCC  
 3228 TCCCCCTTCC CTTTTTTTTC GGTGAGACCT TTTATTTTCT ACTCTTTTCA  
 3278 GAGGGGTTTT CTGTTTGTTC GGGTTTGTTC TCTTGCTGTG ACTGAAACAA  
 3328 GAAGGTTATT GCAGCAAAAA TCAGTAACAA AAAATAGTAA CAATACCTTG  
 3378 CAGAGGAAAG GTGGGAGGAG AGGAAAAAAG GGAAATTTTT AAAGAAATCT

FIG. 1C-3

3428 ATATATTGGG TTGTTTTTTT TTTTGTITTT TGTTTTTTTT TTTTGGGTTT  
 3478 TTTTTTTTTA CTATATATCT TTTTTTTGTT GTCTCTAGCC TGATCAGATA  
 3528 GGAGCACAAG CAGGGGACGG AAAGAGAGAG ACACTCAGGC GGCAGCATTG  
 3578 CCTCCCAGCC ACTGAGCTGT CGTGCCAGCA CCATTCTGG TCACGCAAAA  
 3628 CAGAACCCAG TTAGCAGCAG GGAGACGAGA ACACCACACA AGACATTTTT  
 3678 CTACAGTATT TCAGGTGCCT ACCACACAGG AAACCTTGAA GAAAATCAGT  
 3728 TTCTAGAAGC CGCTGTTACC TCTTGTTTAC AGTTTATATA TATATGATAG  
 3778 ATATGAGATA TATATATAAA AGGTACTGTT AACTACTGTA CAACCCGACT  
 3828 TCATAATGGT GCTTTCAAAC AGCGAGATGA GTAAAAACAT CAGCTTCCAC  
 3878 GTTGCCCTTCT GCGCAAAGGG TTTCACCAAG GATGGAGAAA GGGAGACAGC  
 3928 TTGCAGATGG CCGGTTCTCA CCGTGGGCTC TTCCCCTTGG TTTGTAACGA  
 3978 AGTGAAGGAG GAGAACTTGG GAGCCAGGTT CTCCCTGCCA AAAAGGGGGC  
 4028 TAGATGAGGT GGTGGGCCCC GTGGACAGCT GAGAGTGGGA TTCATCCAGA  
 4078 CTCATGCAAT AACCCCTTGA TTGTTTTCTA AAAGGAGACT CCCTCGGCAA  
 4128 GATGGCAGAG GGTACGGAGT CTTCAGGCCC AGTTTCTCAC TTTAGCCAAT  
 4178 TCGAGGGCTC CTTGTGGTGG GATCAGAACT AATCCAGAGT GTGGGAAAGT  
 4228 GACAGTCAAA ACCCCACCTG GAGCAAATAA AAAACATAC AAAACGTAAA  
 4278 AAAAAAAAAA AAAAAA

FIG. 1C-4

Human DNMT3B1 DNA Sequence:

1 GGCCGCGAAT TCGGCACGAG CCCTGCACGG CCGCCAGCCG GCCTCCCGCC  
51 AGCCAGCCCC GACCCGCGGC TCCGCCGCC AGCCGCGCCC CAGCCAGCCC  
101 TCGGCAGGA AAGCATGAAG GGAGACACCA GGCATCTCAA TGGAGAGGAG  
151 GACGCCGGCG GGAGGAAGA CTCGATCTC GTCAACGGGG CCTGCAGCGA  
201 CCAGTCCTCC GACTCGCCCC CAATCCTGGA GGCTATCCG ACCCCGGAGA  
251 TCAGAGGCCG AAGATCAAGC TCGCGACTCT CCAAGAGGGA GGTGTCCAGT  
301 CTGCTAAGCT ACACACAGGA CTTGACAGGC GATGGCGACG GGAAGATGG  
351 GGATGGCTCT GACACCCAG TCATGCCAAA GCTCTTCCG GAAACCAGGA  
401 CTCGTTGAGA AAGCCCAGCT GTCCGAATC GAAATAACAA CAGTGTCTCC  
451 AGCCGGGAGA GGCACAGGCC TTCCCCACGT TCCACCCGAG GCCGGCAGGG  
501 CCGCAACCAT GTGGACGAGT CCCCCGTGGA GTTCCCGGCT ACCAGGTCCC  
551 TGAGACGGCG GGCAACAGCA TCGGCAGGAA CGCCATGGCC GTCCCCTCCC  
601 AGCTCTTACC TTACCATCGA CCTCACAGAC GACACAGAGG ACACACATGG  
651 GACGCCCCAG AGCAGCAGTA CCCCCTACGC CCGCCTAGCC CAGGACAGCC  
701 AGCAGGGGGG CATGGAGTCC CCGCAGGTGG AGGCAGACAG TGGAGATGGA  
751 GACAGTTCAG AGTATCAGGA TGGGAAGGAG TTTGGAATAG GGGACCTCGT  
801 GTGGGGAAAG ATCAAGGGCT TCTCCTGGTG GCCCGCCATG GTGGTGTCTT  
851 GGAAGGCCAC CTCCAAGCGA CAGGCTATGT CTGGCATGCG GTGGGTCCAG  
901 TGGTTTGGCG ATGGCAAGTT CTCGAGGTC TCTGCAGACA AACTGGTGGC  
951 ACTGGGGCTG TTCAGCCAGC ACTTTAATT GGCCACCTTC AATAAGCTCG  
1001 TCTCCTATCG AAAAGCCATG TACCATGCTC TGGAGAAAGC TAGGGTGCGA  
1051 GCTGGCAAGA CCTTCCCCAG CAGCCCTGGA GACTCATTGG AGGACCAGCT  
1101 GAAGCCCATG TTGGAGTGGG CCCACGGGGG CTTCAAGCCC ACTGGGATCG  
1151 AGGGCCTCAA ACCCAACAAC ACGCAACCAG TGGTTAATAA GTCGAAGGTG

FIG. 1D-1

1201 CGTCGTGCAG GCAGTAGGAA ATTAGAATCA AGGAAATACG AGAACAAGAC  
1251 TCGAAGACGC ACAGCTGACG ACTCAGCCAC CTCTGACTAC TCCCCGCAC  
1301 CCAAGCGCCT CAAGACAAAT TGCTATAACA ACGGCAAAGA CCGAGGGGAT  
1351 GAAGATCAGA GCCGAGAACA AATGGCTTCA GATGTTGCCA ACAACAAGAG  
1401 CAGCCTGGAA GATGGCTGTT TGTCTTGTGG CAGGAAAAAC CCCGTGTCCT  
1451 TCCACCCTCT CTTTGAGGGG GGGCTCTGTC AGACATGCCG GGATCGCTTC  
1501 CTTGAGCTGT TTTACATGTA TGATGACGAT GGCTATCAGT CTTACTGCAC  
1551 TGTGTGCTGC GAGGGCCGAG AGCTGCTGCT TTGCAGCAAC ACGAGCTGCT  
1601 GCCGGTGTTT CTGTGTGGAG TGCCTGGAGG TGCTGGTGGG CACAGGCACA  
1651 GCGGCCGAGG CCAAGCTTCA GGAGCCCTGG AGCTGCTACA TGTGTCTCCC  
1701 GCAGCGCTGT CATGGCGTCC TCGGGCGCCG GAAGGACTGG AACGTGCGCC  
1751 TGCAGGCCTT CTTACCAGT GACACGGGGC TTGAATACGA AGCCCCAAG  
1801 CTGTACCCTG CCATTCCCGC AGCCCGAAGG CGGCCCATTC GAGTCCTGTC  
1851 ATTGTTTGAT GGCATCGCGA CAGGCTACCT AGTCCTCAAA GAGTTGGGCA  
1901 TAAAGGTAGG AAAGTACGTC GCTTCTGAAG TGTGTGAGGA GTCCATTGCT  
1951 GTTGAACCG TGAAGCACGA GGGGAATATC AAATACGTGA ACGACGTGAG  
2001 GAACATCACA AAGAAAAATA TTGAAGAATG GGGCCCATTT GACTTGGTGA  
2051 TTGCCGGAAG CCCATGCAAC GATCTCTCAA ATGTGAATCC AGCCAGGAAA  
2101 GGCCTGTATG AGGTACAGG CCGGCTCTTC TTGGAATTTT ACCACCTGCT  
2151 GAATTACTCA CGCCCCAAGG AGGTGATGA CCGGCCGTTT TTCTGGATGT  
2201 TTGAGAATGT TGTAGCCATG AAGTTGGCG ACAAGAGGGA CATCTCAGG  
2251 TTCCTGGAGT GTAATCCAGT GATGATTGAT GCCATCAAAG TTTCTGCTGC  
2301 TCACAGGGCC CGATACTTCT GGGGCAACCT ACCCGGGATG AACAGCCCCG  
2351 TGATAGCATC AAAGAATGAT AACTCGAGC TGCAGGACTG CTTGGAATAC  
2401 AATAGGATAG CCAAGTTAAA GAAAGTACAG ACAATAACCA CCAAGTCGAA

FIG. 1D-2

2451 CTCGATCAAA CAGGGGAAAA ACCAACTTTT CCCTGTTGTC ATGAATGGCA  
 2501 AAGAAGATGT TTTGTGGTGC ACTGAGCTCG AAAGGATCTT TGGCTTTCCT  
 2551 GTGCACTACA CAGACGTGTC CAACATGGGC CGTGGTGCCC GCCAGAAGCT  
 2601 GCTGGGAAGG TCCTGGAGCG TGCCTGTCAT CCGACACCTC TTCGCCCCCTC  
 2651 TGAAGGACTA CTTTGCATGT GAATAGTTCC AGCCAGGCCC CAAGCCCCT  
 2701 GGGGTGTGTG GCAGAGCCAG GACCCAGGAG GTGTGATTCC TGAAGGCATC  
 2751 CCCAGGCCCT GCTCTTCCTC AGCTGTGTGG GTCATACCGT GTACCTCAGT  
 2801 TCCCTCTTGC TCAGTGGGGG CAGAGCCACC TGA CTCTTGC AGGGGTAGCC  
 2851 TGAGGTGCCG CCTCCTTGTG CACAAATCAG ACCTGGCTGC TTGGAGCAGC  
 2901 CTAACACGGT GCTCATTTTT TCTTCTCCTA AAAC TTAAA ACTTGAAGTA  
 2951 GGTAGCAACG TGGCTTTTTT TTTTCCCTT CCTGGGTCTA CCACTCAGAG  
 3001 AAACAATGGC TAAGATACCA AAACCACAGT GCCGACAGCT CTCCAATACT  
 3051 CAGGTTAATG CTGAAAAATC ATCCAAGACA GTTATTGCAA GAGTTTAATT  
 3101 TTTGAAAAC TGGTACTGCT ATGTGTTTAC AGACGTGTGC AGTTGTAGGC  
 3151 ATGTAGCTAC AGGACATTTT TAAGGGCCCA GGATCGTTTT TTCCAGGGC  
 3201 AAGCAGAAGA GAAAATGTTG TATATGTCTT TTACCCGGCA CATTCCCCTT  
 3251 GCCTAAATAC AAGGGCTGGA GTCTGCACGG GACCTATTAG AGTATTTTCC  
 3301 ACAATGATGA TGATTTACAG AGGGATGACG TCATCATCAC ATTCAGGGCT  
 3351 ATTTTTTCCC CCACAAACCC AAGGGCAGGG GCCACTCTTA GCTAAATCCC  
 3401 TCCCCGTGAC TGCAATAGAA CCCTCTGGGG AGCTCAGGAA GGGGTGTGCT  
 3451 GAGTTCTATA ATATAAGCTG CCATATATTT TGTAGACAAG TATGGCTCCT  
 3501 CCATATCTCC CTCTTCCCTA GGAGAGGAGT GTGAAGCAAG GAGCTTAGAT  
 3551 AAGACACCCC CTCAAACCCA TTCCCTCTCC AGGAGACCTA CCCTCCACAG  
 3601 GCACAGGTCC CCAGATGAGA AGTCTGCTAC CCTCATTTCT CATCTTTTAA  
 3651 CTAAACTCAG AGGCAGTGAC AGCAGTCAGG GACAGACATA CATTCTCAT

FIG. 1D-3

3701 ACCTTCCCA CATCTGAGAG ATGACAGGA AACTGCAA GCTCGGTGCT  
3751 CCCTTTGGAG ATTTTAAAT CCTTTTAT TCCATAAGAA GTCGTTTTA  
3801 GGGAGAACGG GAATTCAGAC AAGCTGCATT TCAGAAATGC TGTATAATG  
3851 GTTTTAACA CCTTTACTC TTCTTACTGG TGCTATTTG TAGAATAAGG  
3901 AACACGTTG ACAAGTTTTG TGGGCTTTT TATACACTTT TTAAATCTC  
3951 AACTTCTAT TTTATGTTT AACGTTTCA TTAAATTTT TTTGTAAGT  
4001 GAGCCACGAC GTAACAAATA TGGGAAAAA ACTGTGCCTT GTTCAACAG  
4051 TTTTGCTAA TTTTAGGCT GAAAGATGAC GGATGCCTAG AGTTTACCTT  
4101 ATGTTTAATT AAAATCAGTA TTTGTCTAAA AAAAAAAAAA AAAAA

FIG. 1D-4

### Mouse Dnmt3a Protein

1 MPSSGPGDTS SSSLEREDDR KEGEEQEENR GKEERQEPSA TARKVGRPGR  
51 KRKHPPVESS DTPKDPAVTT KSQPMQDSG PSDLLPNGDL EKRSEQPPEE  
101 GSPAAGQKGG APAEGEGTET PPEASRAVEN GCCVTKEGRG ASAGEGKEQK  
151 QTNIESMKME GSRGRLRGGL GWESSLRQRP MPRLTFQAGD PYYISKRRKD  
201 EWLARWKREA EKKAKVIAVM NAVEENQASG ESQKVEEASP PAVQQPTDPA  
251 SPTVATTPEP VGGDAGDKNA TKAADDEPEY EDGRGFGIGE LVWGKLRGFS  
301 WWPGRIVSWW MTGRSRAAEG TRWMMFQDG KFSVVCVEKL MPLSSFCSAF  
351 HQATYNQPM YRKAIYEVQ VASSRAGKLF PACHDSDESD SGKAVEVQNK  
401 QMIEWALGGF QPSGPKGLEP PEEENPYKE VYTDMMVEPE AAAYAPPPPA  
451 KKPRKSTTEK PKVKEIIDER TRERLVYEVQ QKCRNIEDIC ISCGSLNVTL  
501 EHPLFIGGMC QNCKNCFLEC AYQYDDGYQ SYCTICCGGR EVLMCGNNNC  
551 CRCFCVECVD LLVGPGAAQA AIKEDPWNCY MCGHKGTYGL LRRREDWPSR  
601 LQMFFANNHD QEFDPPKVYP PVPAEKRKPI RVLSLFDGIA TGLLVLKDLG  
651 IQVDRIASE VCEDSITVGM VRHQGKIMYV GDVRSVTQKH IQEWGPFDLV  
701 IGGSPCNDLS IVNPARKGLY EGTGRLFFEF YRLLDARPK EGDRPFFWL  
751 FENVVAMGVS DKRDISRFLE SNPVMIDAKE VSAHRARYF WGNLPGMNRP  
801 LASTVNDKLE LQECLEHGRI AKFSKVRTIT TRSNSIKQCK DQHFVFMNE  
851 KEDILWCTEM ERVFGFPVHY TDVSNMSRLA RQRLGRSWS VPVIRHLFAP  
901 LKEYFACV\*

FIG. 2A

Mouse Dnmt3b1 Protein

1 MKGDSRHLNE EEGASGYEEC IIVNGNFSDQ SSDTKDAPSP PVLEAICTEP  
51 VCTPETRGRR SSSRLSKREV SSLLNYTQDM TGDGDRDDEV DDGNGSDILM  
101 PKLTRETKDT RTRSESPAVR TRHSNGTSSL ERQRASPRIT RGRQGRHHVQ  
151 EYPVEFPATR SRRRRASSSA STPWSSPASV DFMEEVTPKS VSTPSVDLSQ  
201 DGDQEGMDTT QVDAESRDGD STEYQDDKEF GIGDLWVGKI KGFSWWPAMV  
251 VSWKATSKRQ AMPGMRWVQW FGDGKFSEIS ADKLVALGLF SQHFNLATFN  
301 KLVSYRKAMY HTLEKARVRA CKTFSSSPGE SLEDQLKPML EWAHGCKPT  
351 GIEGLKPNKK QPVVNKSKVR RSDSRNLEPR RRENKSRRT TNSAASESP  
401 PPKRLKTNSY GKGDRGEDEE SRERMASEVT NNKGNLEDRC LSCGKKNPVS  
451 FHPLFEGGLC QSCRDRFLEL FYMYDEGQY SYCTVCCEGR ELLLCNTSC  
501 CRCFCVECLE VLVGAGTAED AKLQEPWSCY MCLPQRCHGV LRRRKDWNMR  
551 LQDFFTTDPD LEEFEPKLY PAIPAAKRRP IRVLSLFDGI ATGYLVKEL  
601 GIKVEKYIAS EVCAESIAVG TVKHEGQIKY VNDVRKITKK NIEEWGPFDL  
651 VIGGSPCNDL SNVNPARKGL YEGTGRLFFE FYHLLNYTRP KEGDNRPFfw  
701 MFENVVAMKV NDKKDISRFL ACNPVMIDAI KVSAAHRARY FWGNLPGMNR  
751 PVMASKNDKL ELQDCLEFSR TAKLKKVQTI TTKSNSIRQG KNQLFPVVMN  
801 GKDDVLWCTE LERIFGFPAH YTDVSNMGRG ARQKLLGRSW SVPVIRHLFA  
851 PLKDYFACE\*

FIG. 2B

### Human DNMT3A Protein

1 MPAMPSSGPG DTSSSAAERE EDRKDGESEQE EPRGKEERQE PSTTARKVGR  
51 PGRKRKHPPV ESGDTPKDPA VISKSPSMAQ DSGASELLPN GDLKLRSEPD  
101 PEEGSPAGGQ KGGAPAELEG AAETLPEASR AVENGCCPK EGRGAPAEAG  
151 KEQKETNIES MKMEGSRGRL RGGLGWESSL RQRPMPLTF QAGDPYYISK  
201 RKRDEWLARW KREAEEKAKV IAGMNAVEEN QGPGESQKVE EASPPAVQQP  
251 TDPASPTVAT TPEPVGSDAG DKNATKAGDD EPEYEDGRGF GIGELVWGKL  
301 RGFSWMPGRI VSWMTGRSR AAEGTRWVMW FGDGKFSVVC VEKLMPLSSF  
351 CSAFHQATYN KQPMYRKAIY EVLQVASSRA GKLPVCHDS DESDTAKAVE  
401 VQNKPMIEWA LGGFQPSGPK GLEPPEEEKN PYKEVYTMW VEPEAAAYAP  
451 PPPAKKPRKS TAEKPKVKEI IDERTRERLV YEVRQKCRNI EDICISGSL  
501 NVTLEHPLFV GGMCQNCKNC FLECAQYDD DGYQSYCTIC CGGREVLMOG  
551 NNNCCRCFCV ECVDLLVGPG AAQAAIKEDP WNCYMGCHKG TYGLLRRED  
601 WPSRLQMFFA NNHDQEFDP KYPVPVPAEK RKPIRVLSLF DGIATGLLVL  
651 KDLGIQVDY IASEVCDSI TVGMVRHQGK IMYVGDVRSV TQKHIEWGP  
701 FDLVIGGSPC NDLSIVNPAR KGLYEGTGRL FFEFYRLLD ARPKEGDDRP  
751 FFWLFENVVA MGVSDKRDIS RFLESNPVM DAKEVSAHR ARYFWGNLPG  
801 MNRPLASTVN DKLELQECLE HGRIAKFSKV RTITTRSNSI KQKGDQHFVP  
851 FMNEKEDILW CTEMERVFGF PVHYTDVSNM SRLARQRLG RSWSPVIRH  
901 LFAPLKEYFA CV\*

FIG. 2C

### Human DNMT3B1 Protein

1 MKGDTRHLNG EEDAGGREDS ILVNGACSDQ SSDSPPILEA IRTPEIRGRR  
51 SSSRLSKREV SSLLSYTQDL TGDGDGEDGD GSDTPVMPKL FRETRTRSES  
101 PAVRTRNNNS VSSRERHRPS PRSTRGRQGR NHVDESPVEF PATRSLRRRA  
151 TASAGTPWPS PPSSYLTI DL TDDTEDTHGT PQSSSTPYAR LAQDSQQGGM  
201 ESPQVEADSG DGDSEYQDG KEFGIGDLVW GKIKGFSWWP AMVVSWKATS  
251 KRQAMSGMRW VQWFGDGKFS EVSADKLVAL GLFSQHFNLA TFNKLVSYRK  
301 AMYHALEKAR VRAGKTFPSS PGDSLEDQLK PMLEWAHGGF KPTGIEGLKP  
351 NNTQPVVNS KVRRA GSRKL ESRKYENKTR RRTADDSATS DYCPAPKRLK  
401 TNCYNNGKDR GDEDQSREQM ASDVANNKSS LEDGCLSCGR KNPVSFHPLF  
451 EGGLECQTCRD RFLELFMYD DDGYQSYCTV CCEGRELLLC SNTSCCRFC  
501 VECLEVLVGT GTAAEAKLQE PWSCYMCLPQ RCHGVLRRRK DWNVRLQAFF  
551 TSDTGLEYEA PKLYPAIPAA RRRPIRVLSL FDGIATGYLV LKELGIKVGK  
601 YVASEVCEES IAVGTVKHEG NIKYVNDVRN ITKKNIEEWG PFDLVIGGSP  
651 CNDLSNVNPA RKGLYEGTGR LFFEFYHLLN YSRPKEGDDR PFFWMFENVV  
701 AMKVGDKRDI SRFLECNPVM IDAIKVSAAH RARYFWGNLP GMNRPVIASK  
751 NDKLELQDCL EYNRIAKLKK VQTITTKSNS IKQGKNQLFP VMNGKEDVL  
801 WCTELERIFG FPVHYTDVSN MGRGARQKLL GRSWSVPVIR HLFAPLKDYF  
851 ACE\*

FIG. 2D

Dnmt3a 1 MPSSGPGDTSSSSLEREDDRKEGEEQEENRGKEERQEPSATARKVGRPGR 50  
Dnmt3a 51 KRKHPPVESSDTPKDPVTTKSQPM AQDSGPSD....LLPNGDLEKRSEP 96  
Dnmt3b 1 .....MKGDSRHLNEEEGASGYEECIIVNGNFSDQSSD 33  
Dnmt3a 97 QPEEGSP....AAGQKGGAPAEGETETPPEAS.RAVENGCCVTKE..GR 139  
Dnmt3b 34 TKDAPSPPVLEAICTEPVCTPETRGRSSSRLSKREVSSLLNYTQDMTGD 83  
Dnmt3a 140 G.....ASAGEG.....KEQKQTNIESMKMEGSRGRLRGGLGWESSLRQ 178  
Dnmt3b 84 GDRDDEVDDGNGSDILMPKLTRETKDTRRSESPAVRTRHSNGTSSLERQ 133  
Dnmt3a 179 RPMPRLTFQAGDPYYISKRRDEWLARWKREA EKKAKVIAVMNAVEENQA 228  
Dnmt3b 134 RASPRITRGRQGRHHV.....QEYPVEFPATRSRRRRASSASTPWSSPA 178  
Dnmt3a 229 SGESQKVEEASPPAVQQPTDPASPTVATTPEPVGGDAGDKNATKAADDEP 278  
Dnmt3b 179 SVDF..MEEVTPKSVSTP....SVDLSQDGDQEGMDTTQVDAESRDGDST 222  
Dnmt3a 279 EYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGTRWVMWFG 328  
Dnmt3b 223 EYQDDKEFGIGDLVWGKIKGFSWWPAMVVSWKATSKRQAMPGMRWVQWFG 272  
Dnmt3a 329 DGKFSVVCVEKLMPLSSFCSAFHQATYNKQPMYRKAIYEV LQVASSRAGK 378  
Dnmt3b 273 DGKFSEISADKLVALGLFSQHFNLATFNKLVS YRKAMYHTLEKARVRAGK 322  
Dnmt3a 379 LFPACHDSDES DSGKAVEVQNQMIEWALGGFQPSGPKGLEPPEEEK..N 426  
Dnmt3b 323 TF.....SSSPGESLEDQLKPMLEWAHGGFKPTGIEGLKPNKKQPVVN 365  
Dnmt3a 427 PYKEVYTMW.VEP.....EAAAYAPPPPAKKPRKSTTEKPK 462  
Dnmt3b 366 KSKVRRSDSRNLEPRRRENKSRRRTTND SAASESPPPKRLKTN SYGGKDR 415

FIG.3A-1

Dnmt3a	463	VKEIIDERTRERLVYEVQRKCRNIEDICISCGSLNVTLEHPFFIGGMCQN	512
		.    .   :     :    :            :	
Dnmt3b	416	GE...DEESRERMASEVTNNKGNLEDRCLSCGKKNPVSFHPLFEGGLCQS	462
Dnmt3a	513	CKNCFLECAQQYDDDGYSYCTICCGGREVLNCGNNCCRCFCVECVDLL	562
		: .         :     :      .  :   .         . :.	
Dnmt3b	463	CRDRFLELFYMYDEDEGYSYCTVCCEGRELLCSNTSCCRCFCVECLEVL	512
Dnmt3a	563	VGPGAAQAAIKEDPWNCYMGHKGTYGLLRRREDWPSRLQMFFANNHD.Q	611
		:   :   .     . : .    .           .   :	
Dnmt3b	513	VGAGTAEDAKLQEPWSCYMCLPQRCHGVLRRRKDOWNMRLQDFFTTDPDLE	562
Dnmt3a	612	EFDPPKVYPVPVPAEKRPVIRVLSLFDGIATGLLVKDLGIQVDRIASEV	661
		:   .    :     :                :   .  :	
Dnmt3b	563	EFEPKLYPAIPAARRPVIRVLSLFDGIATGYLVKELGIKVEKYIASEV	612
Dnmt3a	662	CEDSITVGMVRHQGKIMYVGDVRSVTQKHIEWGPFDLVIGGSPCNDLSI	711
		:       : .          : .  :	
Dnmt3b	613	CAESIAVGTVKHEGQIKYVNDVRKITKKNIEEWGPFDLVIGGSPCNDLSN	662
Dnmt3a	712	VNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWL FENVVAMGVSD	761
		.      .      :        .	
Dnmt3b	663	VNPARKGLYEGTGRLFFEFYHLLNYTRPKEGDNRPFFWMFENVVAMKVND	712
Dnmt3a	762	KRDISRFLESNPVMIDAKEVSAHRARYFWGNLPGMNRPLASTVNDKLEL	811
		:             .                   . ..	
Dnmt3b	713	KKDISRFLACNPVMIDAIVSAHRARYFWGNLPGMNRPMASKNDKLEL	762
Dnmt3a	812	QECLEHGRIAKFSKVRTITTRSNSIKQKQDQHPVFMNEKEDILWCTEME	861
		:            .     :     :    .           : :     :	
Dnmt3b	763	QDCLEFSRTAKLKKVQTITTKSNSIROGKNQLFPVVMNGKDDVLWCTELE	812
Dnmt3a	862	RVFGFPVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACV*	909
		:                 :                  :	
Dnmt3b	813	RIFGFPAHYTDVSNMGRGARQKLLGRSWSVPVIRHLFAPLKDYFACE*	860

FIG.3A-2

DNMT3A 1 MPAMPSSGPGDTSSSAAEREEDRKDGEEQEEPRGKEERQEPSTTARKVGR

DNMT3A 51 PGRKRKHPPVESGDTPKDPAVISKSPSMAQDSGASELLPNGDLEKRSEPO

DNMT3B 1 .....MKGDTRHLNGEEDAGGREDSILVNGACSDQSSDSP

DNMT3A 101 PEEGSPAGGQKGGAPAEGEGAAETLPEASRAVENGCCTPKEGRGAPAEAG

DNMT3B 36 PILEAIRTPEIRGGWASSRLSKREVSSLLSYTQDLTGDDGDGEDGDSPT

DNMT3A 151 KEQKETNIESMKMEGSRGRLRGGLGWESSLRQRPMPRLTFQAGDPYYISK

DNMT3B 86 VMPKLFRETRTRSESPAVRTRNNNSVSSRERHRPSRSTRGRQGRNHVDE

DNMT3A 201 RKRDEWLARWKREAEKKAKVIAGMNAVEENQGPGESQKVVEASPPAVQQP

DNMT3B 136 SPVEFPATRSLRRRATASAGTPWPSPPSSYLTDLTDDTDTH..GTPQS

DNMT3A 251 TDPASPTVATTPEPVGSDAGDKNATKAGDDEPEYEDGRGFGIGELVWGKL

DNMT3B 184 SSTPYARLAQDSQQGGMES PQVEADSGDGSSEYQDGKEFGIGDLVWGKI

DNMT3A 301 RGFSWWPGRIVSWWMTGRSRAAEGTRWVMWFGDGKFSVVCVEKLMPLSSF

DNMT3B 234 KGFSWWPAMVVSWKATSKRQAMSGMRWVQWFGDGKFSEVSADKLVALGLF

DNMT3A 351 CSAFHQATYNKQPMYRKAIYEV LQVASSRAGKLPVCHDSDESDTAKAVE

DNMT3B 284 SQHFNLATFNKLVS YRKAMYHALEKARVRAGKTFP.....SSPGDSLE

DNMT3A 401 VQNKPMIEWALGGFQPSGPKGLEP....PEEEKNPYKEVYTDMWVE....

DNMT3B 327 DQLKPMLEWAHGGFKPTGIEGLKPNNTQPVVNKS KVR RAGSRKLESRKYE

DNMT3A 443 .....PEAAAYAPPPAKKPRKSTAEKPKVKEIIDERTRERLVYEVQR

DNMT3B 377 NKTRRRRTADDSATSDYCPAPKRLKTNCYNNGKDRGDEDQSREQMASDVAN

FIG.3B-1



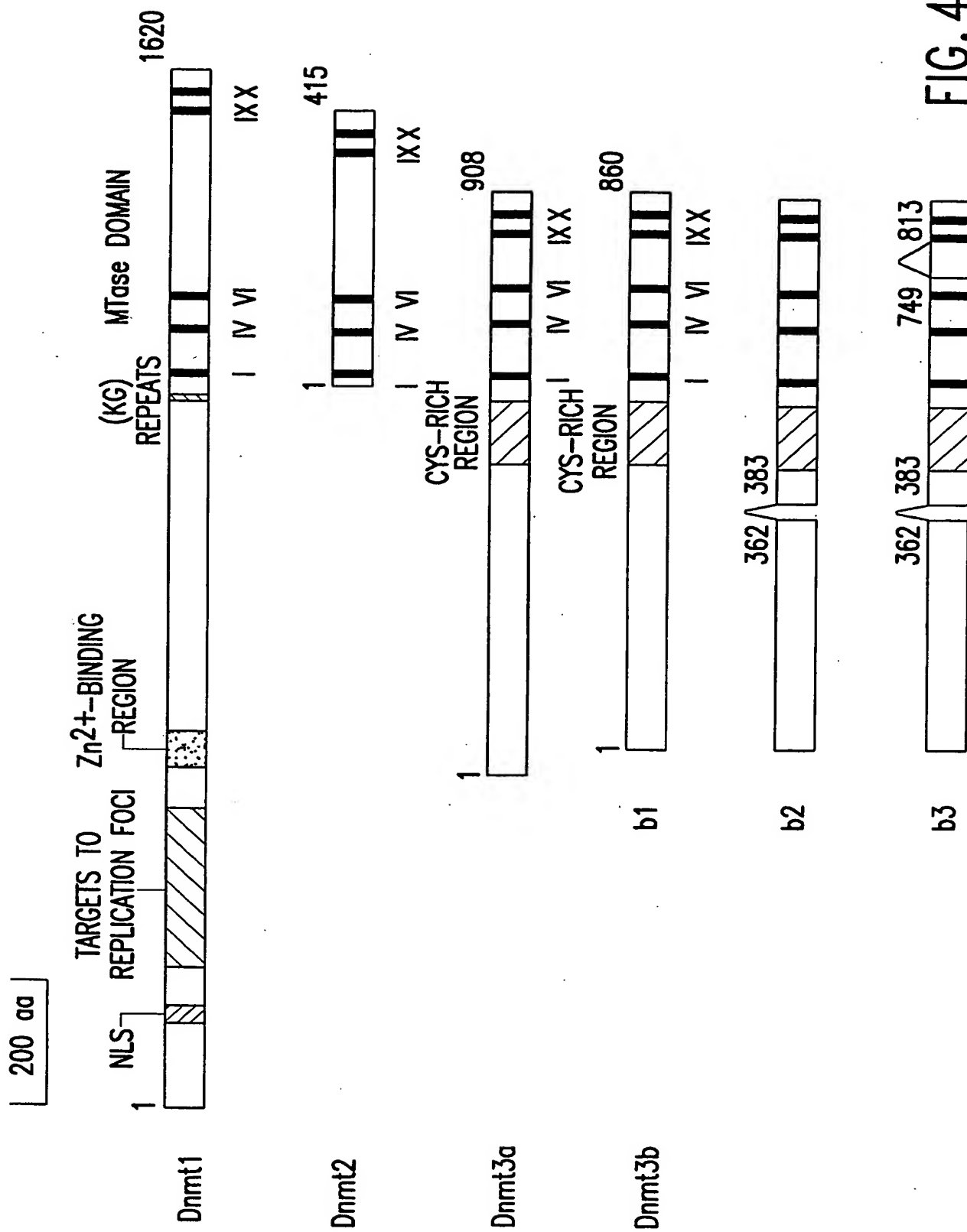


FIG. 4A

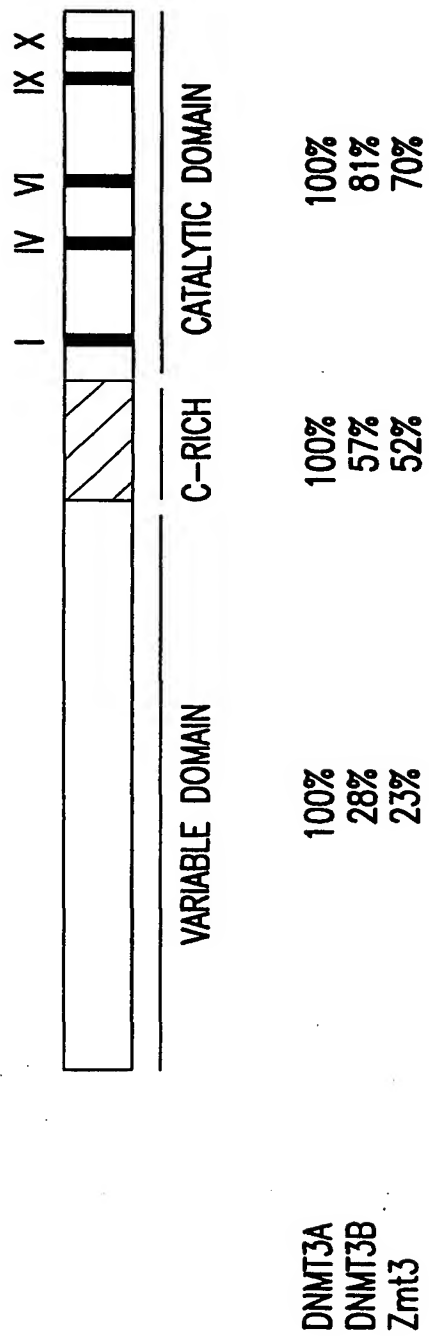
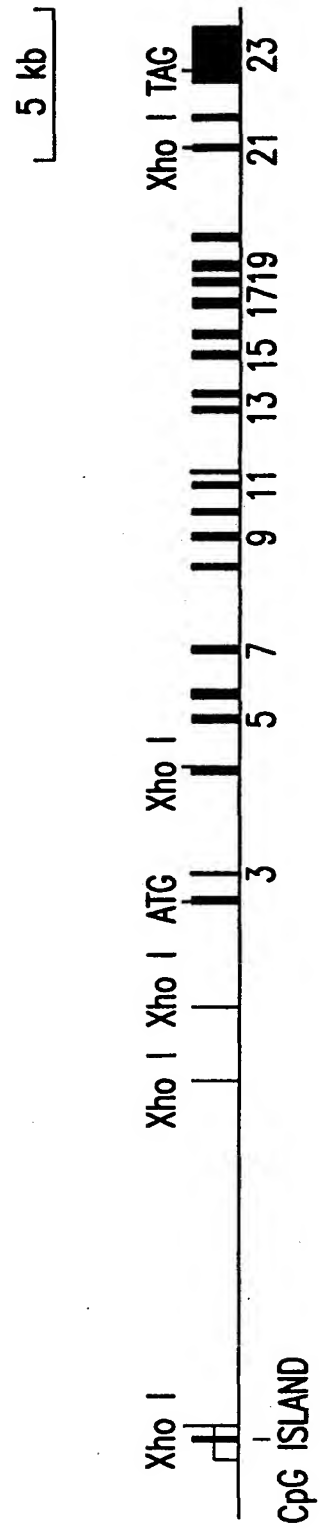


FIG. 4B



**FIG. 4C**

Exon1 (>=90bp) CGGCAGgtgagcgcgcccggg.intron(17618bp).tggtttctccacagGAAAGC  
 Exon2 ( 148bp) TCAGAGgtggctgggcagtg.intron( 887bp).CTGTTTCTCTACAGGCCGAA  
 Exon3 ( 62bp) ACACAGgtatggtctctctc.intron( 3343bp).tggttccttataaaagGACTTG  
 Exon4 ( 102bp) CCAGCTgtaagtagccacacc.intron( 1642bp).ctctctgcttctagTCCGA  
 Exon5 ( 125bp) ACCAGgttggtcccccagatg.intron( 602bp).tcctctgtccacagTCCCTG  
 Exon6 ( 222bp) TATCAGgtatggccgagaggg.intron( 1403bp).tggttttcttccagGATGGG  
 Exon7 ( 159bp) TCCGAGgtgagtcggggaag.intron( 2588bp).gtcttctcttttagGTCTCT  
 Exon8 ( 108bp) CTGGAGgtaacatgggatgag.intron( 917bp).actctgcctttgcagAAAGCT  
 Exon9 ( 145bp) AACCAggtggaatgagtccc.intron( 765bp).ttttccctcaaaagTGGTTA  
 Exon10 ( 60bp) AATACGgtatttccttctgt.intron( 1813bp).aattacctttcacagAGAACAA  
 Exon11 ( 126bp) GCCGAGgtgattgttggttac.intron( 115bp).ttcttttctcaatagAACAAA  
 Exon12 ( 45bp) TGAAGgtaacgttctctccc.intron( 1095bp).ctgttttcttctacagATGGCT  
 Exon13 ( 80bp) TGCCGGgtaaagtccctcctact.intron( 417bp).ctctctggctgccagGATCGC  
 Exon14 ( 113bp) CTGCCGgtgagcactgggccc.intron( 1160bp).tgccactgggtccagGTGTTT  
 Exon15 ( 184bp) GAATACgtaaagccacaggctc.intron( 600bp).ttccttacctggcagGAAGCC  
 Exon16 ( 85bp) CGACAGgtgagttcggggaac.intron( 824bp).ctctggcccccacagGCTACC  
 Exon17 ( 146bp) AAAATgtgagggcagtcctgt.intron( 536bp).gtctctctctttcagATTGAA  
 Exon18 ( 91bp) TGATGgtgagcatccttctc.intron( 352bp).cttttctgagcacagAGGGTA  
 Exon19 ( 149bp) CTGGAGgtgaggaatctggg.intron( 958bp).tctttctccccacagTGTAAAT  
 Exon20 ( 86bp) GAACAGgtaaacaaaggctct.intron( 2867bp).tttggctgttcccagGCCCGT  
 Exon21 ( 70bp) GCCAAGttaaagaaagtacag.intron( 801bp).cattttgttctccagTTAAAG  
 Exon22 ( 119bp) CGAAAGgtgagcaaggctgca.intron( 1434bp).ctccggtacccccagGATCTT  
 Exon23 (1585bp)

FIG.4D

	I	IV	VI
DNMT1	DVFSGCGGLSEGFHQAG	DVEMLCGGPPCQGFSGMNR	YRPRFFLLNNRNFSFKR
Dnmt1	DVFSGCGGLSEGFHQAG	DVEMLCGGPPCQGFSGMNR	YRPRFFLLNNRNFSYRR
MET1 (Ath)	DIFAGCGGLSHGLKKAG	QVDFINGGPPCQGFSGMNR	FRPRYFLLNNVRTFVSFNK
Masc1	DTFCGGGVSIGARQAG	HVDILHLSPPCQTFSRAHT	VRPRLFTVEETDGIHDRQS
Masc2	DIFAGCGGLTLGLDLSG	EVDFIYGGPPCQGFSGVNR	YKPRFVLLNNVKGLITTKL
Dnmt2	ELYSGIGGMHHALRESH	SFNMIILMSPPCQPFTRIGL	KLPKYILLNNWKGFEVSST
M. Spr	SLFSGIGAFEAALRNIG	EFDLLVGGSPCQSFSAVGH	KQPKFFVFENNKGGLINHDK
DNMT3A	SLFDGIATGLLVLDLG	PFDLVIIGSPCNDLSIVNP	DRPFFWLFEINWAMGVSDK
Dnmt3a	SLFDGIATGLLVLDLG	PFDLVIIGSPCNDLSIVNP	DRPFFWLFEINWAMGVSDK
DNMT3B	SLFDGIATGYLVVKELG	PFDLVIIGSPCNDLSNVNP	DRPFFMFEINWAMKVGDK
Dnmt3b	SLFDGIATGYLVVKELG	PFDLVIIGSPCNDLSNVNP	NRPFFMFEINWAMKVNDK
Zmt3	SLFDGIATGYLVLRDLG	PFDLLIGSPCNDLSIVNP	PQPFFWLFEINWTFMQTHVK
consensus	--F-G-----G	-----GG-PC--S-N-	--P-F--ENN-----

	IX	X
DNMT1	RWVSVRECARSQGFP	LFGNILDKRVQGNVAVPPPLAKAIG
Dnmt1	RWVSVRECARSQGFP	FFGNILDRHRQVGNVAVPPPLAKAIG
MET1 (Ath)	RILTVRECARSQGFP	FAGNINHHRQIGNAVPPPLAFALG
Masc1	RKFTVRELACIQGFP	FVGTLTDKRRIIGNAVPPPLSAAIM
Masc2	RVYTVRELARAQGFP	GLGGVKKWHRNIGNAVVPVPLGEQIG
Dnmt2	RYFTPKEIANLQGFP	EKTTVKQRYRLGNLSLNVHVAKLL
M. Spr	RRLTPLECFRLOAFD	AGISNSQLYKQTGNSITVTVLESIF
DNMT3A	DILWCTEMERVFGFP	SNMSRLARQRLIGRSWSVPVIRHLF
Dnmt3a	DILWCTEMERVFGFP	SNMSRLARQRLIGRSWSVPVIRHLF
DNMT3B	DVLWCTELERIFGFP	SNMGRGARQKLLIGRSWSVPVIRHLF
Dnmt3b	DVLWCTELERIFGFP	SNMGRGARQKLLIGRSWSVPVIRHLF
Zmt3	DHIWITELEKIFGFP	KSMGRPORQVRVLGKSWSPVIRHLL
consensus	-----E--R--GFP	-----R--G-----P-----

FIG. 5A

DNMT3A	EDICISCG.....SLNVTLLEHPLFVGGMCQNCNKCFLECAQYQDDDGYSYCT
Dnmt3a	EDICISCG.....SLNVTLLEHPFFIGGMCQNCNKCFLECAQYQDDDGYSYCT
DNMT3B	EDGLSCG.....RKNPVSFHPLFEGLCQTCDRFLELFMYWDDDGYSYCT
Dnmt3b	EDRLSCG.....KKNPVSFHPLFEGLCQCSDRFLELFMYWDEDGYQS YCT
Znt3	EDFLSCG.....SMSVDI IHPLFEGLCTNCKNFETLTRYDEDDGYQS YCT
ATRX_Human	IVSCTAGQQWNHFQKDS IYRHPS LQLICKNCFKYMSDDISRDS DGMDEQCR
ATRX_Mouse	IVSCTAGQQWNHFQKDS IYRHPS LQLICKNCFKYMSDDISRDS DGMDEQCR
Consensus	C C C C C C C C
DNMT3A	ICCGGREVL MCGNNCCRCFCVECDLLVGPGAQAAIKE.DPWNCVMCGHKGT
Dnmt3a	ICCGGREVL MCGNNCCRCFCVECDLLVGPGAQA AIKE.DPWNCVMCGHKGT
DNMT3B	VCCEGRELLCSNTSCRCFCVECLEVLVTGTAAEAKLQ.EPWS CYMCLPQR
Dnmt3b	VCCEGRELLCSNTSCRCFCVECLEVL VGAGTAEADAKLQ.EPWS CYMCLPQR
Znt3	VCCSGMEVILCAHDCSRSCFDCLD ILVCQGTFDRCLKNV.DPWTCYLCAPE TS
ATRX_Human	WCAEGGNLIC..DFCHNAFCKKCILRNLRGRKELSTIMDENNQWYCICHPE PL
ATRX_Mouse	WCAEGGNLIC..DFCHNAFCKKCILRNLRGRKELSTIMDENNQWYCICHPE PL
Consensus	C C C C C C C C

**FIG. 5B**

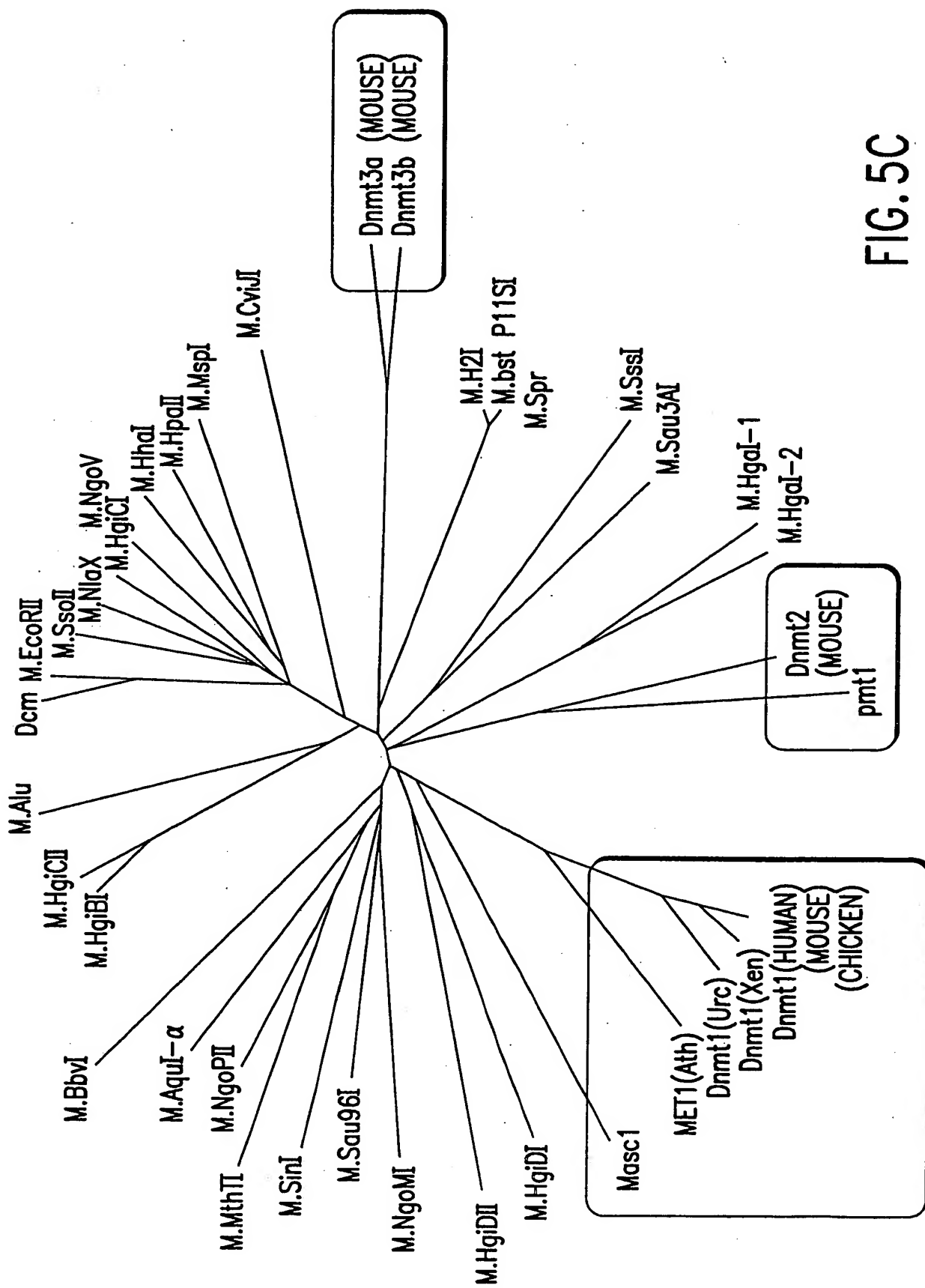


FIG. 5C

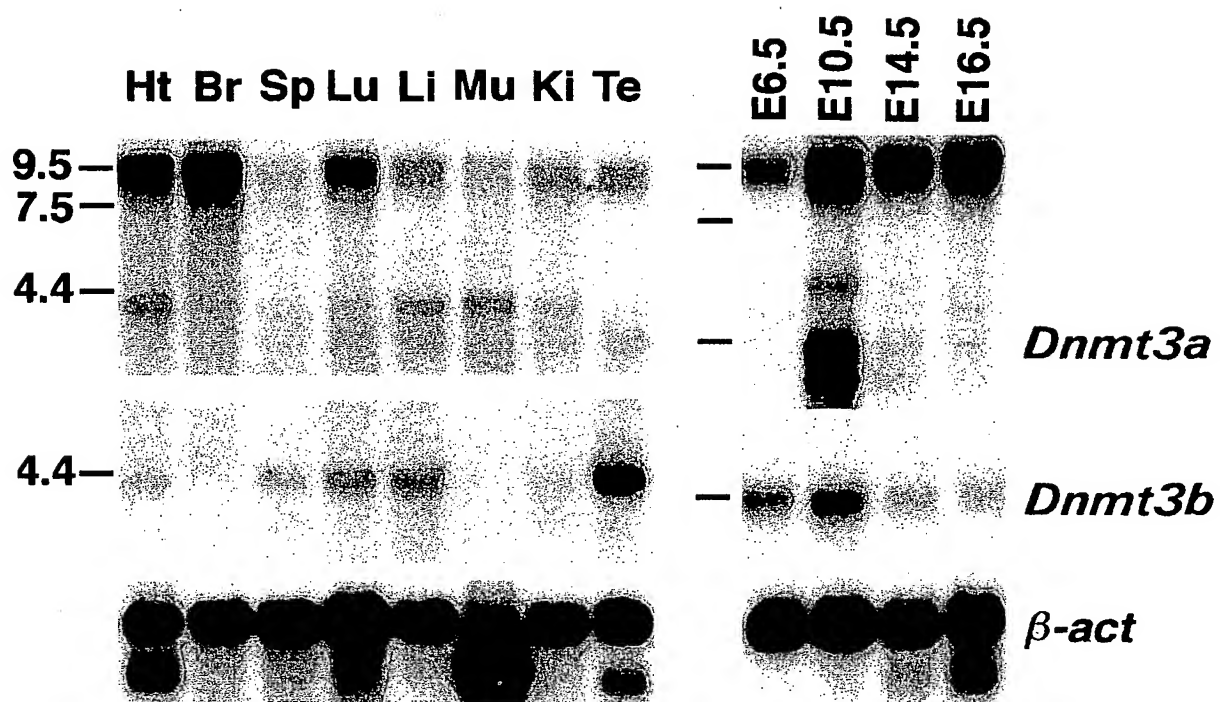


FIG.6A

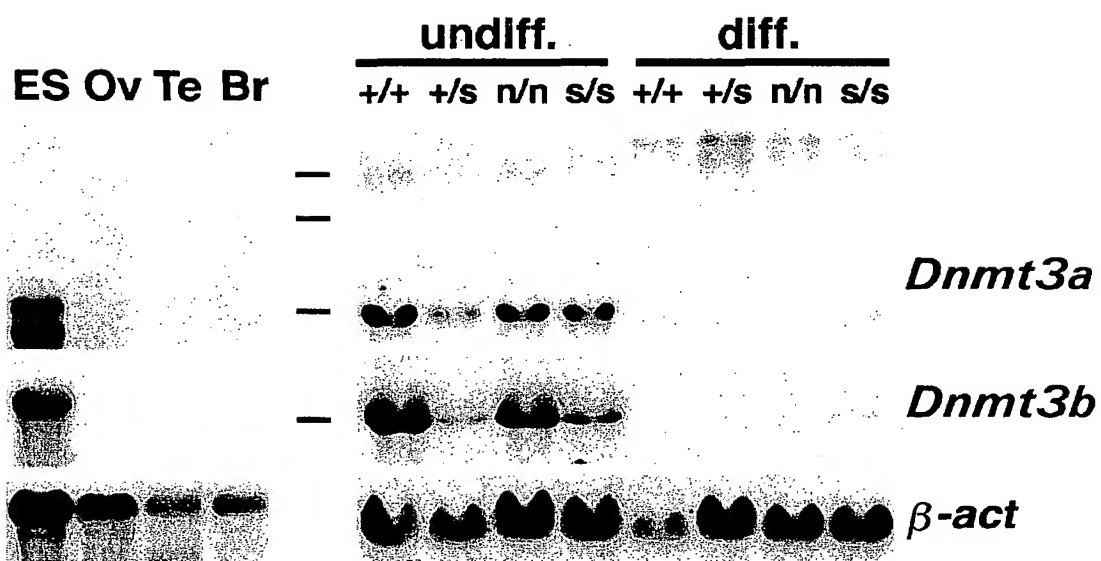


FIG.6B

FIG.6C

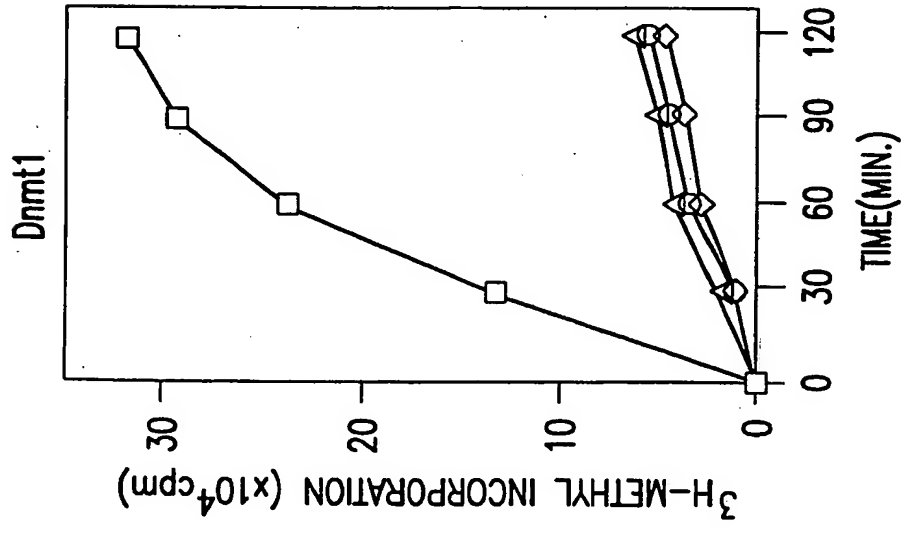


FIG. 7A

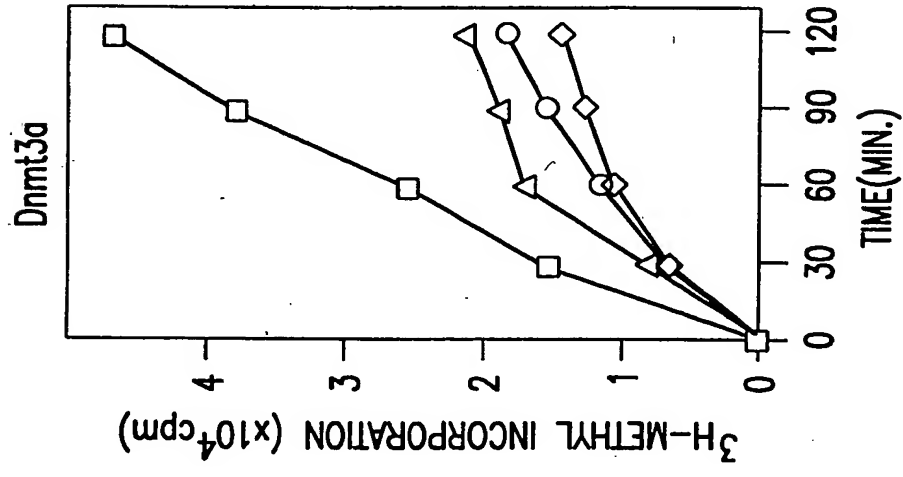


FIG. 7B

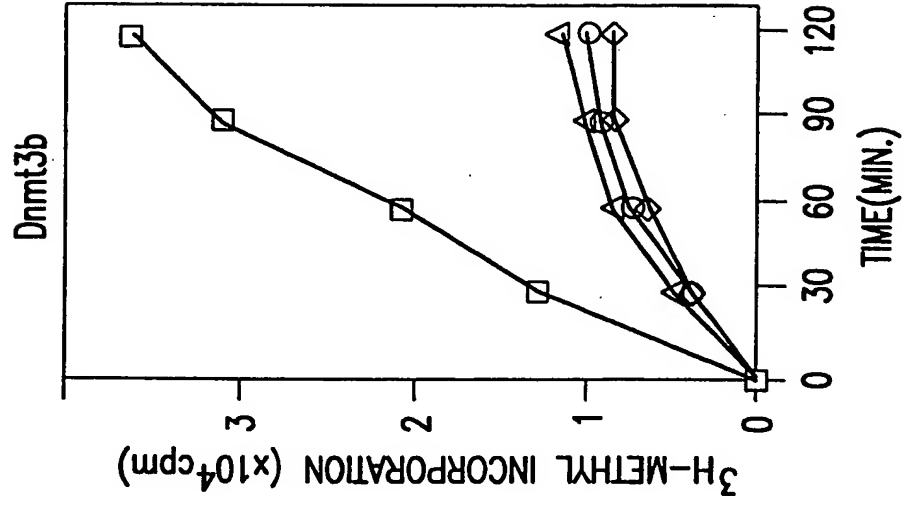


FIG. 7C



FIG.7D

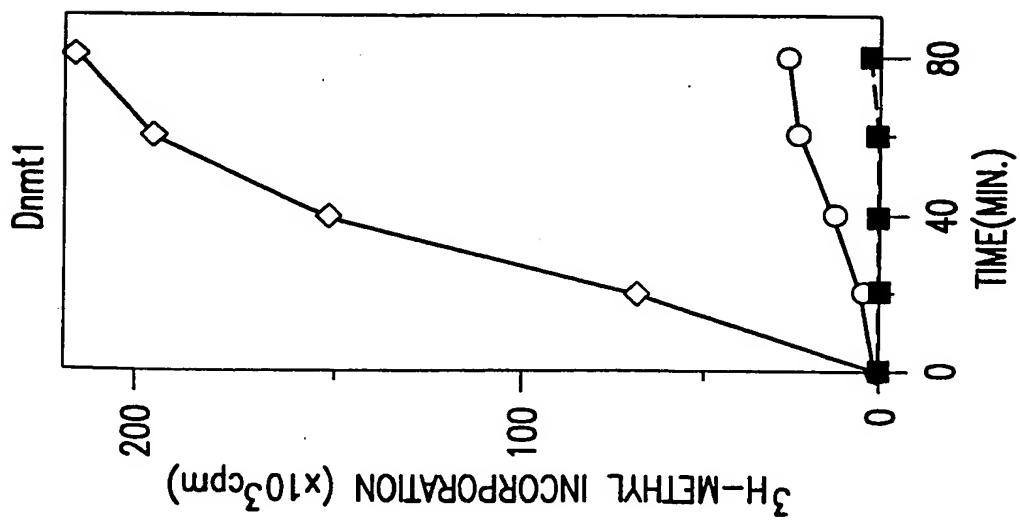


FIG. 8A

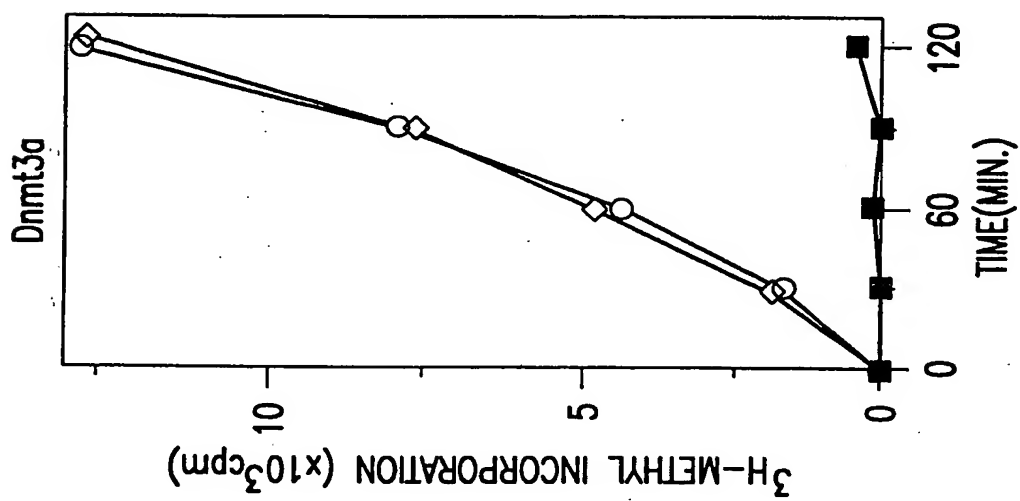


FIG. 8B

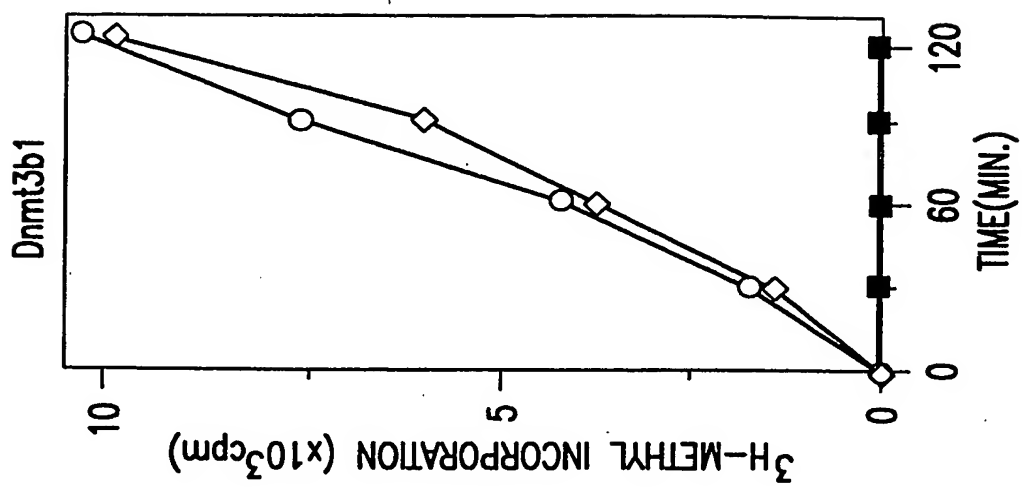


FIG. 8C

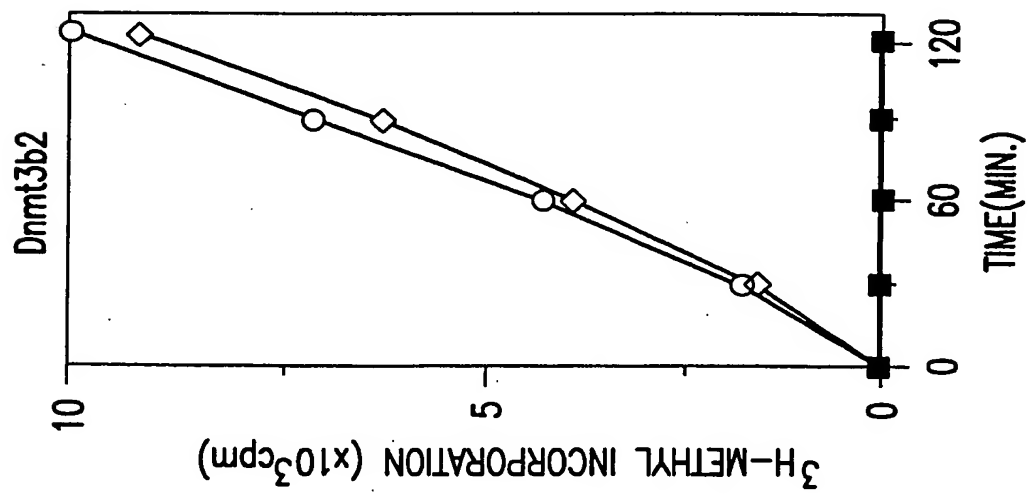


FIG. 8D

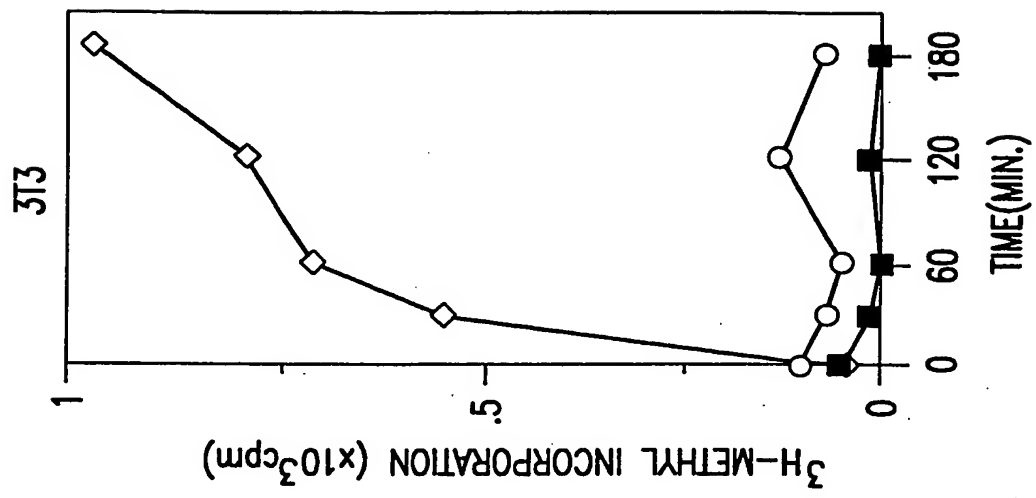


FIG. 8E

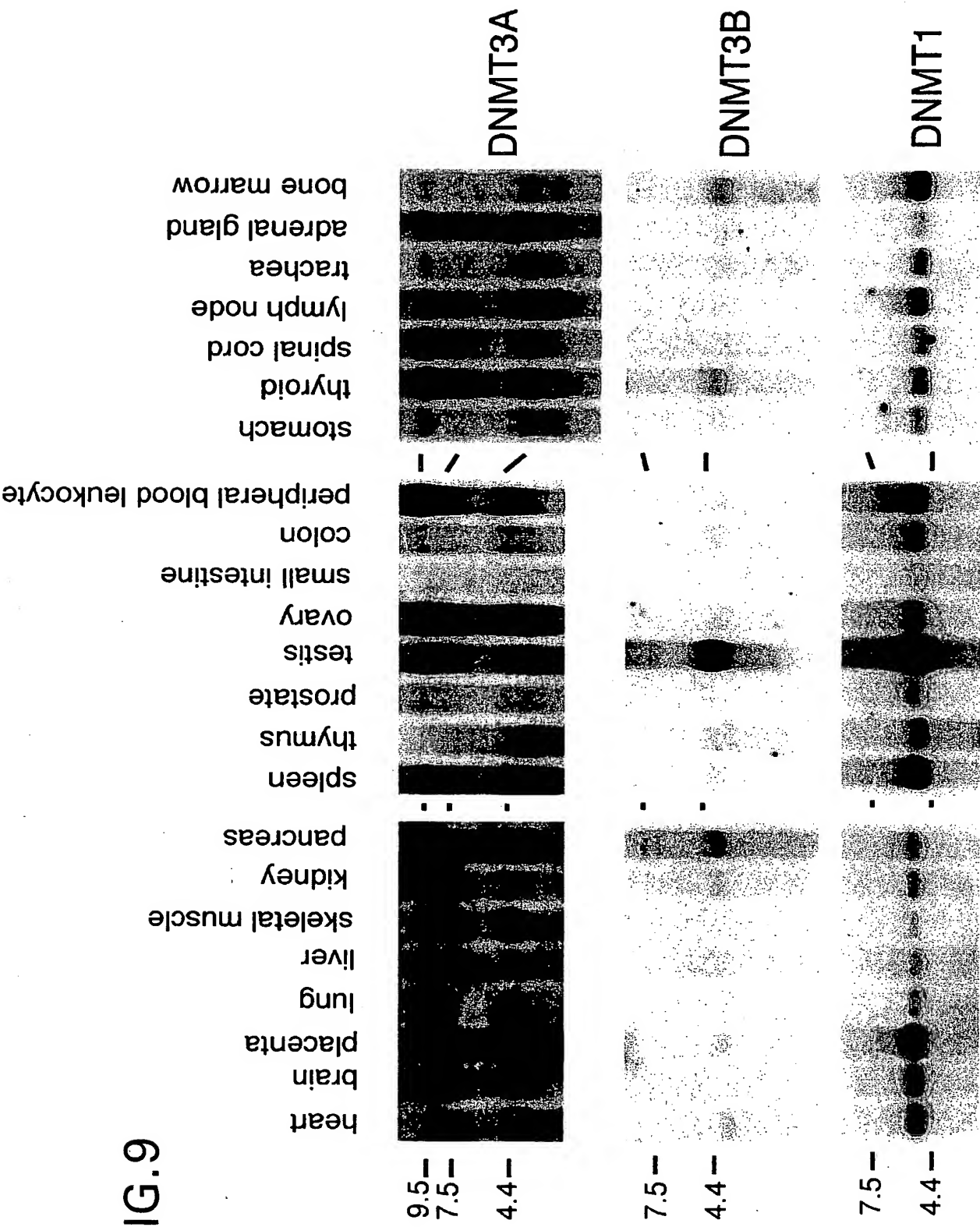


FIG.9

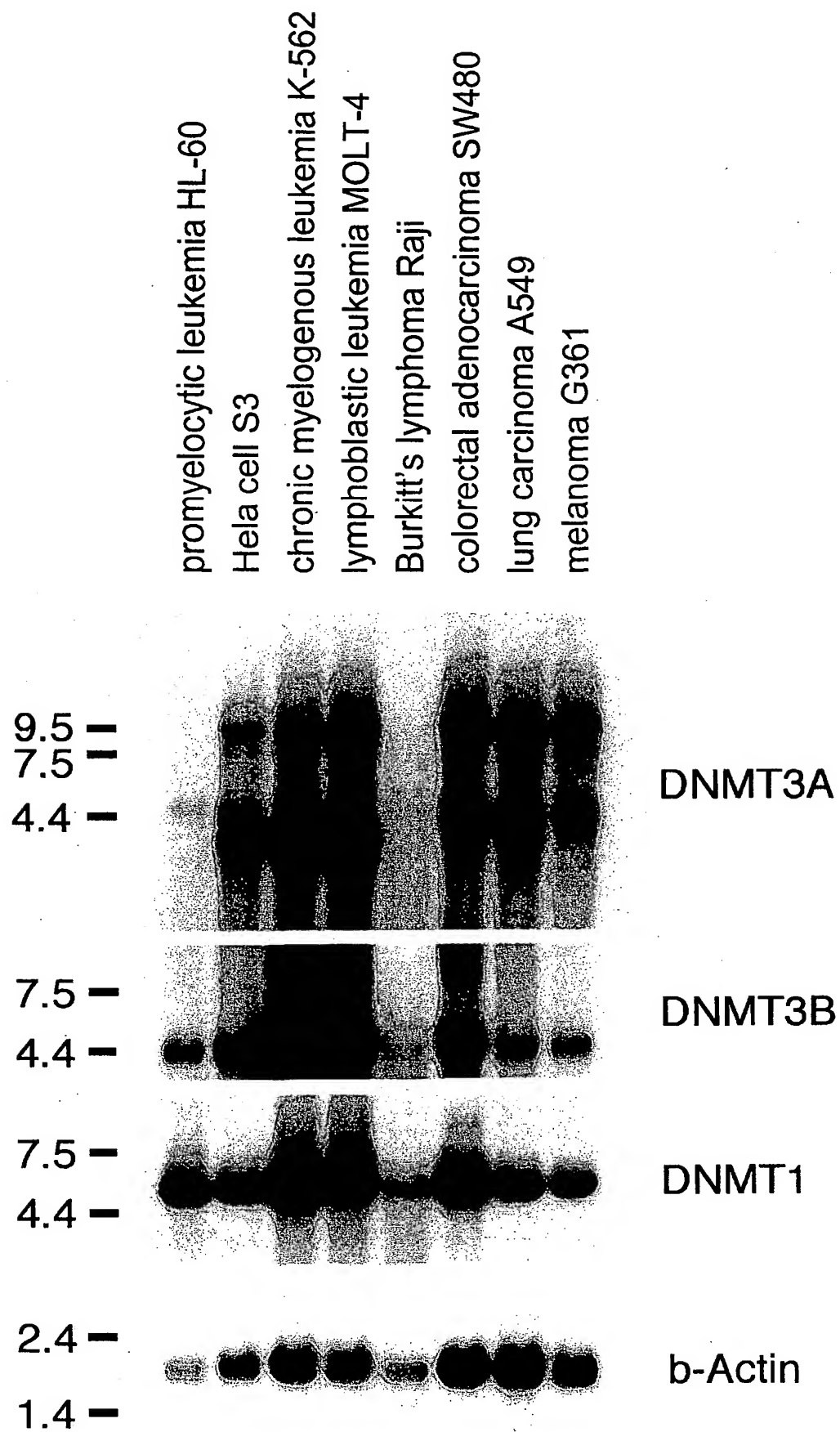
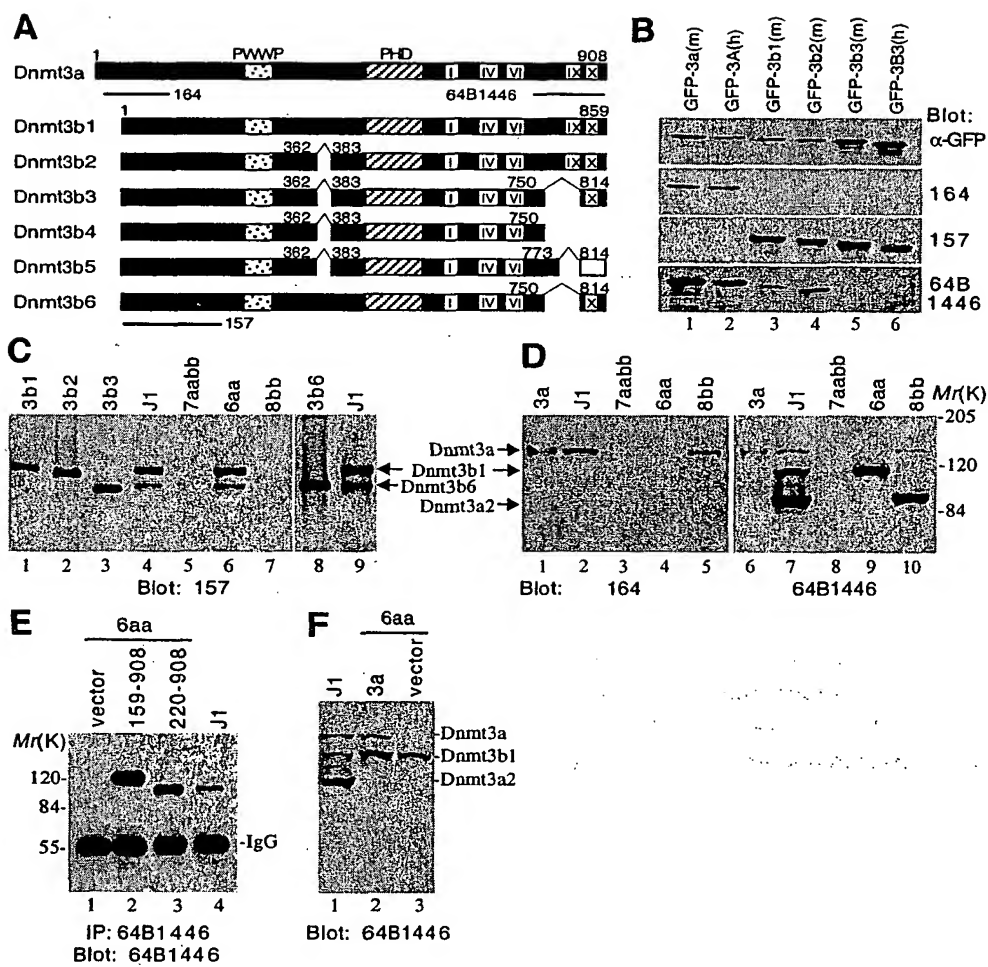
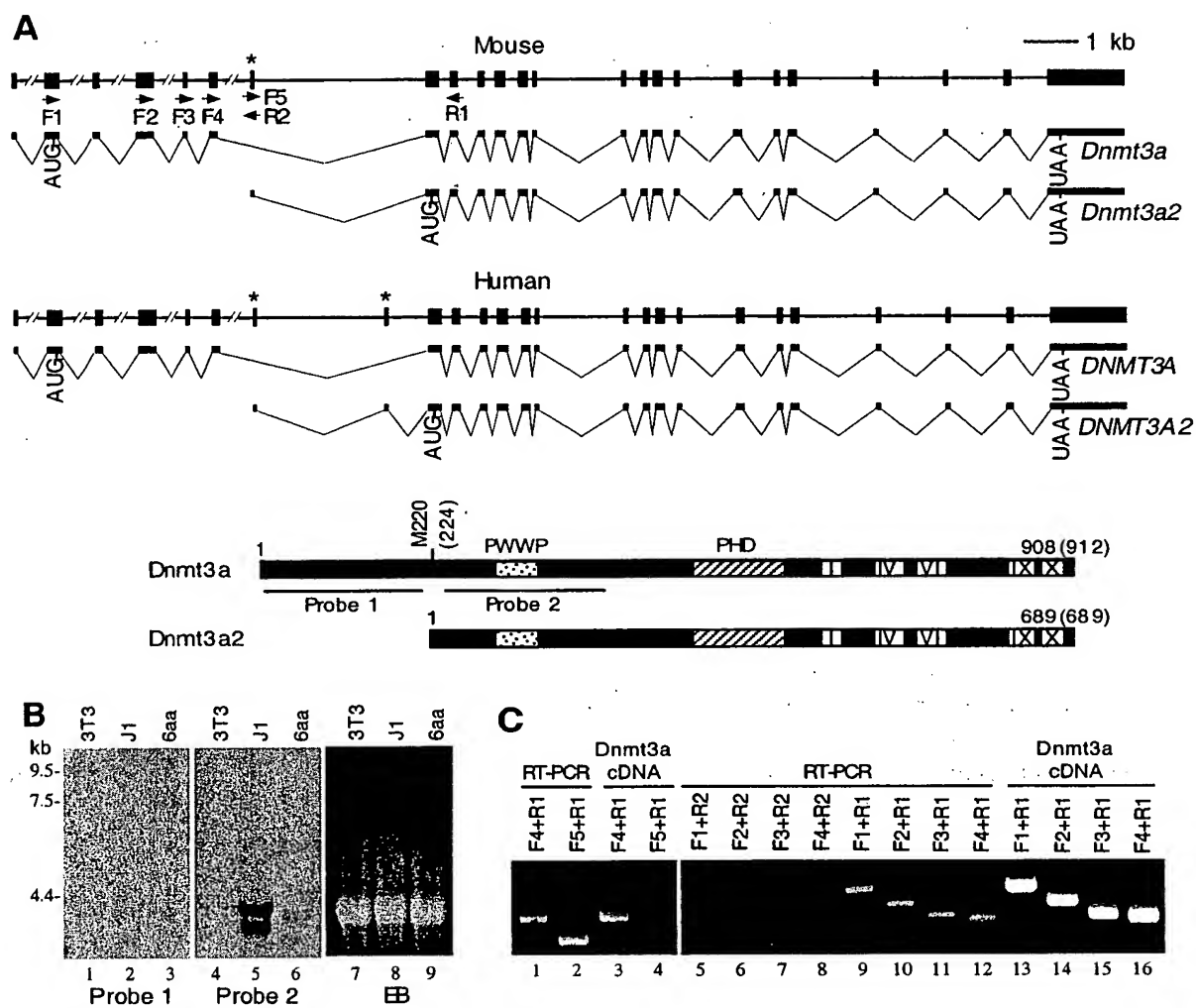


FIG.10



**FIG. 11**



**FIG. 12**

Mouse Dnmt3a2 cDNA sequence:

1 cgcggcccaaaccccaacgccccctgcccctccccccagacggggcagctatttacagagc  
60 ttcggggccggggctcacacctgagctgtactgcagagggggctgcacctggccttatggg  
119 ctgagaagaaagccaaggtaatgtcagtaataatgctgtggaagagaaccaggcctct  
178 ggagagtctcagaaggtggaggaggccagccctcctgctgtgcagcagcccacggaccc  
237 tgcttctccgactgtggccaccacccctgagccagtaggaggggatgctggggacaaga  
296 atgctaccaaagcagccgacgatgagcctgagtatgaggatggccggggccttggcatt  
355 ggagagctggtgtgggggaaacttcggggccttctcctggtggccaggccgaattgtgtc  
414 ttggtggatgacaggccggagccgagcagctgaaggcactcgctgggtcatgtggttcg  
473 gagatggcaagtctcagtggtgtgtgtggagaagctcatgccgctgagctccttctgc  
532 agtgcattccaccaggccacctacaacaagcagcccatgtaccgcaaagccatctacga  
591 agtcctccagggtggccagcagccgtgccgggaagctgtttccagcttgccatgacagtg  
650 atgaaagtgcagtggaaggctgtggaagtgcagaacaagcagatgattgaatggggcc  
709 ctcggtggcttcagccctcggttcctaagggcctggagccaccagaagaagagaagaa  
768 tccttacaaggaagtttacaccgacatgtgggtggagcctgaagcagctgcttacgccc  
827 cccccccaccagccaagaaacccagaaagagcacaaacagagaaacctaagggtcaaggag  
886 atcattgatgagcgcacaagggagcggctggtgtatgaggtgcgccagaagtgcagaaa  
945 catcgaggacatttgtatctcatgtgggagcctcaatgtcacctggagcacccactct  
1004 tcattggaggcatgtgccagaactgtaagaactgcttcttgagtggtgcttaccagtat  
1063 gacgacgatgggtaccagtcctattgcaccatctgctgtggggggcggtgaagtgtcat  
1122 gtgtgggaacaacaactgctgcagggtgcttttgtgtcgagtgtgtggatctcttggtgg  
1181 ggccaggagctgctcaggcagccattaaggaagaccctggaactgctacatgtgcggg  
1240 cataagggcacctatgggctgctgcgaagacgggaagactggccttctcgactccagat  
1299 gttctttgccaataaccatgaccaggaatttgaccccccaaagggtttaccacactgtgc  
1358 cagctgagaagaggaagcccatccgcgtgctgtctctctttgatgggatgtctacaggg  
1417 ctcctggtgctgaaggacctgggcatccaagtggaccgctacattgcctccgaggtgtg  
1476 tgaggactccatcacggtgggcatggtgcggcaccagggaagatcatgtacgtcgggg  
1535 acgtccgcagcgtcacacagaagcatatccaggagtggggcccattcgacctggtgatt  
1594 ggaggcagtccttgcaatgacctctccattgtcaaccctgcccgaagggaactttatga  
1653 gggactggccgcctcttctttgagttctaccgcctcctgcatgatgcgcggcccaagg  
1712 agggagatgatcgcccttcttctggctctttgagaatgtggtggccatgggcttagt  
1771 gacaagaggggacatctcgcgatttcttgagtctaaccctgtgatgatgacgccaaaga  
1830 agtgtctgctgcacacagggcccggttacttctggggtaaccttctggcatgaacaggc  
1889 ctttggcatccactgtgaatgataagctggagctgcaagagtgtctggagcacggcaga  
1948 atagccaagttcagcaaagtgaggaccattaccaccaggtcaaactctataaagcaggg  
2007 caaagaccagcatttccccgtcttcatgaacgagaaggaggacatcctgtggtgcactg  
2066 aaatggaaaggggtgtttggcttccccgtccactacacagacgtctccaacatgagccgc  
2125 ttggcgaggcagagactgctgggcccgatcgaggagcgtgccggtcatccgccacctctt  
2184 cgctccgctgaaggaatattttgcttgtgtgtaagggacatgggggcaaactgaagtag  
2243 tgatgataaaaaagttaaacaaacaaacaaacaaaaacaaaaacaaataaaaacac  
2302 caagaacgagaaaaaaa

FIG. 13A

Mouse Dnmt3a2 amino acid sequence:

1 MNAVEENQASGESQKVEEASPPAVQQPTDPASPTVATTPEPVGGDAGDKNATKAADDEP  
60 EYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGTRWVMWFGDGKFSVVCV  
119 EKL MPLSSFCSAFHQATYNKQPMYRKAIYEVLQVASSRAGKLF PACHDSDES DSGKAVE  
178 VQNKQMI EWALGGFQPSGPKGLEPPEEEKNPYKEVY TDMWVEPEAAAYAPPPPAKKPRK  
237 STTEKPKVKEI IDERTRERLVYEVRQKCRNIEDI CISCGLNVTLEHPLFIGGMCQNCK  
296 NCFLECA YQYDDDGYQSYCTICCGGREVL MCGNNNCCRCFCVECVDLLVGP GAAQAAIK  
355 EDPWNCY MCGHKGT YGLLRREDWPSRLQMFFANNHDQEFDP PKVYPPVPAEK RKPIRV  
414 LSLFDGIATGLLV LKDLGIQVDRIASEVCEDSITVGMVRHQ GKIMYVGDVRSVTQKHI  
473 QEWGPFDLVI GGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPK EGDDRPF FFWL  
532 FENVVAMGVSDK RDISRFLESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKL  
591 ELQECLEHGRI AKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDI LWCTEMERVFGFPV  
650 HYTDVSNMSRLARQ RLLGRSWSVPVIRHLFAPLKEYFACV

**FIG. 13B**

Human DNMT3A2 cDNA sequence:

1 ccgccccccagccccatcgcccccttcccctcccccaagacggggcagctacttccagagc  
60 ttcagggcccgggctcacacctgagcgcgactgcagaggggctgcacctggccttatgg  
119 ggatcctggagcgggttgtgagaaggaatgggcgcgatcgtagcctgaaagacgag  
178 tgtgatacggctgagaagaaagccaaggtcattgcaggaatgaatgctgtggaagaaaa  
237 ccagggggcccggggagtcctcagaaggtggaggaggccagccctcctgctgtgcagcagc  
296 ccactgaccccgcatccccactgtggctaccacgcctgagcccggtgggggtccgatgct  
355 ggggacaagaatgccaccaagcaggcgcgatgacgagccagagtacgaggacggccgggg  
414 ctttggcattggggagctggtgtgggggaaactgcggggcttctcctgggtggccaggcc  
473 gcattgtgtcttgggtggatgacgggcccggagccgagcagctgaaggcaccgcctgggtc  
532 atgtgggttcggagacggcaaattctcagtgggtgtgtgttgagaagctgatgccgctgag  
591 ctcgttttgcagtgcgttccaccaggccacgtacaacaagcagcccatgtaccgcaaag  
650 ccatctacgaggtcctgcaggtggccagcagccgcgcggggaagctgttcccgggtgtgc  
709 cacgacagcgcgatgagagtgcactgcccaaggccgctggagggtgcagaacaagcccatgat  
768 tgaatgggcccctggggggcttccagccttctggccctaagggcctggagccaccagaag  
827 aagagaagaatccctacaaagaagtgtacacggacatgtgggtggaacctgaggcagct  
886 gcctacgcaccacctccaccagccaaaaagccccggaagagcacagcggagaagcccaa  
945 ggtcaaggagattattgatgagcgcacaagagagcggctggtgtacgaggtgcggcaga  
1004 agtgccggaacattgaggacatctgcattctcctgtgggagcctcaatgttaccctggaa  
1063 caccctctctcgttggaggaatgtgccaaaactgcaagaactgcttctctggagtgctgc  
1122 gtaccagtacgacgacgacggctaccagtccactgcaccatctgctgtgggggcccgtg  
1181 aggtgctcatgtgcggaaacaacaactgctgcaggtgcttttgcgtggagtgtgtggac  
1240 ctcttgggtggggccgggggctgccaggcagccattaaggaagacccttggaactgcta  
1299 catgtgcgggcacaaagggtacctacgggctgctgcggcggcgagaggactggccctccc  
1358 ggctccagatgttcttcgctaataaccacgaccaggaatttgaccctccaaagggtttac  
1417 ccacctgtcccagctgagaagaggaagcccatccgggtgctgtctctcttcttgatggaat  
1476 cgctacagggctcctggtgctgaaggacttgggcattcaggtggaccgctacattgcct  
1535 cggaggtgtgtgaggactccatcacggtgggcatggtgcggcaccaggggaagatcatg  
1594 tacgtcggggacgtccgcagcgtcacacagaagcatatccaggagtggggcccattcga  
1653 tctggtgattgggggcagtcacctgcaatgacctctccatcgtcaaccctgctcgcaagg  
1712 gcctctacgagggcactggccggctcttctttgagttctaccgcctcctgcatgatgcg  
1771 cggcccaaggaggaggagatgatcgcccttcttctggctctttgagaatgtggtggccat  
1830 gggcggttagtgacaagaggggacatctcgcgatttctcgagtcacaacctgtgatgatg  
1889 atgccaaagaagtgtcagctgcacacaggggcccgcctacttctggggtaaccttcccgggt  
1948 atgaacaggccggtggcatccactgtgaatgataagctggagctgcaggagtgtctgga  
2007 gcatggcaggatagccaagttcagcaaagtgaggaccattactacgaggtcaaactcca  
2066 taaagcagggcacaagaccagcatttctcgtcttcatgaatgagaaagaggacatctta  
2125 tgggtgcaactgaaatggaaagggtatttgggttcccagtcactatactgacgtctccaa  
2184 catgagccgcttggcgaggcagagactgctgggcccgtcatggagcgtgccagtcattcc  
2243 gccacctcttcgctccgctgaaggagtattttgcgtgtgtgtaagggacatggggggcaa  
2302 actgaggtagcgacacaaaagttaaacaaacaaacaaaaaacacaaaacataataaaaca  
2361 ccaagaacatg

FIG. 13C

Human DNMT3A2 amino acid sequence:

1 MNAVEENQGPGESQKVEEASPPAVQQPTDPASPTVATTPEPVGSDAGDKNATKAGDDEP  
60 EYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGTRWVMWFGDGKFSVVCV  
119 EKL MPLSSFCSAFHQATYNKQPMYRKAIYEV LQVASSRAGKLFVPVCHDSDES DTAKAVE  
178 VQNKPMIEWALGGFQPSGPKGLEPPEEEKNPYKEVY TDMWVEPEAAAYAPPPPAKKPRK  
237 STAEKPKVKEIIDERTRERLVYEVRQKCRNIEDICISCGSLNVTLEHPLFVGGM CQNCK  
296 NCFLECA YQYDDDGYQSYCTICCGGREVL MCGNNNCCRCFCVECVDLLVGP GAAQAAIK  
355 EDPWNCYMC GHKGT YG LLRRREDWPSRLQMFFANNHDQEFDP PKVYPPVPAEKRKPIRV  
414 LSLFDGIATG LLVLKDLGIQVDRIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHI  
473 QEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWL  
532 FENVVAMGVSDKRDISRFLESNPVMIDAKEVSAHRARYFWGNLPGMNRPLASTVNDKL  
591 ELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPV  
650 HYTDVSNMSRLARQRL LGRSWSVPVIRHLFAPLKEYFACV

**FIG. 13D**

		10	20	30	40	50	
Dnmt3a2	1	ccgcccccaa	ccccaacgcc	ccctgcccct	ccccccagac	gggcagctat	50
DNMT3A2	1	ccgcccccag	ccccatcgcc	cccttcccct	cccccaagac	gggcagctac	50
		60	70	80	90	100	
Dnmt3a2	51	ttacagagct	tc-gggccgg	ggctcacacc	tgagctgtac	tgacagagggg	100
DNMT3A2	51	ttccagagct	tcaggggccgc	ggctcacacc	tgagcgcgac	tgacagagggg	100
		110	120	130	140	150	
Dnmt3a2	101	ctgcacctgg	ccttatgg--	-----	-----	-----	150
DNMT3A2	101	ctgcacctgg	ccttatgggg	atcctgggagc	gggttgtgag	aaggaatggg	150
		160	170	180	190	200	
Dnmt3a2	151	-----	-----	-----	-----gctg	agaagaaagc	200
DNMT3A2	151	cgcggtggatc	gtagcctgaa	agacgagtgt	gatacggctg	agaagaaagc	200
		210	220	230	240	250	
Dnmt3a2	201	caaggtatt	gcagtaata	atgctgtgga	agagaaccag	gcctctggag	250
DNMT3A2	201	caaggtcatt	gcaggaata	atgctgtgga	agaaaaccag	gggcccgggg	250
		260	270	280	290	300	
Dnmt3a2	251	agtctcagaa	ggtggaggag	gccagccctc	ctgctgtgca	gcagcccacg	300
DNMT3A2	251	agtctcagaa	ggtggaggag	gccagccctc	ctgctgtgca	gcagcccact	300
		310	320	330	340	350	
Dnmt3a2	301	gaccctgctt	ctccgactgt	ggccaccacc	cctgagccag	taggagggga	350
DNMT3A2	301	gaccccgcat	ccccactgt	ggctaccacg	cctgagcccg	tgggggtccga	350
		360	370	380	390	400	
Dnmt3a2	351	tgctggggac	aagaatgcta	ccaaagcagc	cgacgatgag	cctgagtatg	400
DNMT3A2	351	tgctggggac	aagaatgcc	ccaaagcagg	cgatgacgag	ccagagtacg	400
		410	420	430	440	450	
Dnmt3a2	401	aggatggccg	gggctttggc	attggagagc	tggtgtgggg	gaaacttcgg	450
DNMT3A2	401	aggacggccg	gggctttggc	attggggagc	tggtgtgggg	gaaactgcgg	450
		460	470	480	490	500	
Dnmt3a2	451	ggcttctcct	ggtggccagg	ccgaattgtg	tcttggtgga	tgacaggccg	500
DNMT3A2	451	ggcttctcct	ggtggccagg	ccgcattgtg	tcttggtgga	tgacggggccg	500
		510	520	530	540	550	
Dnmt3a2	501	gagccgagca	gctgaaggca	ctcgctgggt	catgtggttc	ggagatggca	550
DNMT3A2	501	gagccgagca	gctgaaggca	cccgtgggt	catgtggttc	ggagacggca	550
		560	570	580	590	600	
Dnmt3a2	551	agttctcagt	ggtgtgtgtg	gagaagctca	tgccgctgag	ctccttctgc	600
DNMT3A2	551	aattctcagt	ggtgtgtgtt	gagaagctga	tgccgctgag	ctcgttttgc	600
		610	620	630	640	650	
Dnmt3a2	601	agtgcattcc	accaggccac	ctacaacaag	cagcccatgt	accgcaaagc	650
DNMT3A2	601	agtgcgttcc	accaggccac	gtacaacaag	cagcccatgt	accgcaaagc	650
		660	670	680	690	700	
Dnmt3a2	651	catctacgaa	gtcctccagg	tggccagcag	ccgtgccggg	aagctgtttc	700
DNMT3A2	651	catctacgag	gtcctgcagg	tggccagcag	ccgcgcgggg	aagctgttcc	700

**FIG. 13E-1**

Dnmt3a2	701	710	720	730	740	750
DNMT3A2	701	cagcttgcca	tgacagtgat	gaaagtgaca	gtggcaaggc	tgtggaagtg 750
Dnmt3a2	751	760	770	780	790	800
DNMT3A2	751	cagaacaagc	agatgattga	atggggccctc	ggtggcttcc	agccctcggg 800
Dnmt3a2	801	810	820	830	840	850
DNMT3A2	801	ccctaagggc	ctggagccac	cagaagaaga	gaagaatcct	tacaaggaag 850
Dnmt3a2	851	860	870	880	890	900
DNMT3A2	851	tttacaccga	catgtgggtg	gagcctgaag	cagctgctta	cgccccaccc 900
Dnmt3a2	901	910	920	930	940	950
DNMT3A2	901	ccaccagcca	agaaacccag	aaagagcaca	acagagaaac	ctaaggtcaa 950
Dnmt3a2	951	960	970	980	990	1000
DNMT3A2	951	ggagatcatt	gatgagcgca	caagggagcg	gctggtgtat	gaggtgcgcc 1000
Dnmt3a2	1001	1010	1020	1030	1040	1050
DNMT3A2	1001	agaagtgcag	aaacatcgag	gacatttgta	tctcatgtgg	gagcctcaat 1050
Dnmt3a2	1051	1060	1070	1080	1090	1100
DNMT3A2	1051	gtcaccctgg	agcaccact	cttcattgga	ggcatgtgcc	agaactgtaa 1100
Dnmt3a2	1101	1110	1120	1130	1140	1150
DNMT3A2	1101	gaactgcttc	ttggagtgtg	cttaccagta	tgacgacgat	gggtaccagt 1150
Dnmt3a2	1151	1160	1170	1180	1190	1200
DNMT3A2	1151	cctattgcac	catctgctgt	ggggggcggtg	aagtgctcat	gtgtgggaac 1200
Dnmt3a2	1201	1210	1220	1230	1240	1250
DNMT3A2	1201	aacaactgct	gcaggtgctt	ttgtgtcgag	tgtgtggatc	tcttggtggg 1250
Dnmt3a2	1251	1260	1270	1280	1290	1300
DNMT3A2	1251	gccaggagct	gctcaggcag	ccattaagga	agaccctgg	aactgctaca 1300
Dnmt3a2	1301	1310	1320	1330	1340	1350
DNMT3A2	1301	tgtgcgggca	taagggcacc	tatgggctgc	tgcaagacg	ggaagactgg 1350
Dnmt3a2	1351	1360	1370	1380	1390	1400
DNMT3A2	1351	ccttctcgac	tccagatgtt	ctttgccaat	aaccatgacc	aggaatttga 1400
Dnmt3a2	1351	ccctcccggc	tccagatgtt	cttcgcta	aaccacgacc	aggaatttga 1400

**FIG. 13E-2**

		1410	1420	1430	1440	1450	
Dnmt3a2	1401	ccccccaaag	gtttaccac	ctgtgccagc	tgagaagagg	aagccccatcc	1450
DNMT3A2	1401	ccctccaaag	gtttaccac	ctgtcccagc	tgagaagagg	aagccccatcc	1450
		1460	1470	1480	1490	1500	
Dnmt3a2	1451	gcgtgctgtc	tctctttgat	gggattgcta	cagggctcct	ggtgctgaag	1500
DNMT3A2	1451	gggtgctgtc	tctctttgat	ggaatcgcta	cagggctcct	ggtgctgaag	1500
		1510	1520	1530	1540	1550	
Dnmt3a2	1501	gacctgggca	tccaagtggg	ccgctacatt	gcctccgagg	tgtgtgagga	1550
DNMT3A2	1501	gacttgggca	ttcaggtggg	ccgctacatt	gcctcggagg	tgtgtgagga	1550
		1560	1570	1580	1590	1600	
Dnmt3a2	1551	ctccatcacg	gtgggcatgg	tgcggcacca	gggaaagatc	atgtacgtcg	1600
DNMT3A2	1551	ctccatcacg	gtgggcatgg	tgcggcacca	ggggaagatc	atgtacgtcg	1600
		1610	1620	1630	1640	1650	
Dnmt3a2	1601	gggacgtccg	cagcgtcaca	cagaagcata	tccaggagtg	gggccccattc	1650
DNMT3A2	1601	gggacgtccg	cagcgtcaca	cagaagcata	tccaggagtg	gggccccattc	1650
		1660	1670	1680	1690	1700	
Dnmt3a2	1651	gacctggtga	ttggaggcag	tccctgcaat	gacctctcca	ttgtcaaccc	1700
DNMT3A2	1651	gatctggtga	ttgggggcag	tccctgcaat	gacctctcca	tcgtcaaccc	1700
		1710	1720	1730	1740	1750	
Dnmt3a2	1701	tgcccgcgaag	ggactttatg	aggggtactgg	ccgcctcttc	tttgagttct	1750
DNMT3A2	1701	tgctcgcaag	ggcctctacg	agggcactgg	ccggtctctc	tttgagttct	1750
		1760	1770	1780	1790	1800	
Dnmt3a2	1751	accgcctcct	gcatgatgcg	cggcccaagg	aggagatga	tcgccccttc	1800
DNMT3A2	1751	accgcctcct	gcatgatgcg	cggcccaagg	aggagatga	tcgccccttc	1800
		1810	1820	1830	1840	1850	
Dnmt3a2	1801	ttctggctct	ttgagaatgt	ggtggccatg	ggcgtagtg	acaagaggga	1850
DNMT3A2	1801	ttctggctct	ttgagaatgt	ggtggccatg	ggcgtagtg	acaagaggga	1850
		1860	1870	1880	1890	1900	
Dnmt3a2	1851	catctcgcga	tttcttgagt	ctaaccctgt	gatgattgac	gccaagaag	1900
DNMT3A2	1851	catctcgcga	tttctcgagt	ccaaccctgt	gatgattgat	gccaagaag	1900
		1910	1920	1930	1940	1950	
Dnmt3a2	1901	tgtctgctgc	acacagggcc	cgttacttct	ggggtaacct	tcctggcatg	1950
DNMT3A2	1901	tgtcagctgc	acacagggcc	cgctacttct	ggggtaacct	tcccggtatg	1950
		1960	1970	1980	1990	2000	
Dnmt3a2	1951	aacaggcctt	tggcatccac	tgtgaatgat	aagctggagc	tgcaagagtg	2000
DNMT3A2	1951	aacaggccgt	tggcatccac	tgtgaatgat	aagctggagc	tgcaagagtg	2000
		2010	2020	2030	2040	2050	
Dnmt3a2	2001	tctggagcac	ggcagaatag	ccaagttcag	caaagtgagg	accattacca	2050
DNMT3A2	2001	tctggagcat	ggcaggatag	ccaagttcag	caaagtgagg	accattacta	2050
		2060	2070	2080	2090	2100	
Dnmt3a2	2051	ccaggtcaaa	ctctataaag	cagggcaaag	accagcattt	ccccgtcttc	2100
DNMT3A2	2051	cgaggtcaaa	ctccataaag	cagggcaaag	accagcattt	tcctgtcttc	2100

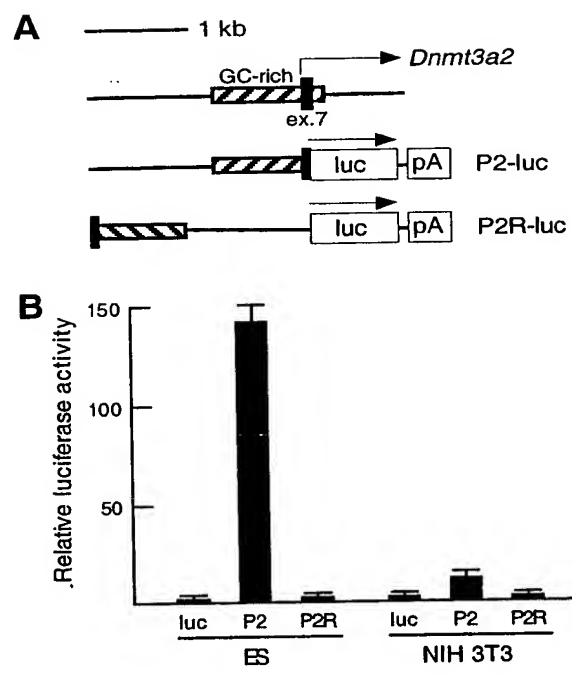
**FIG. 13E-3**

		2110	2120	2130	2140	2150	
Dnmt3a2	2101	atgaacgaga	aggaggacat	cctgtggtgc	actgaaatgg	aaaggggtgtt	2150
DNMT3A2	2101	atgaatgaga	aagaggacat	cttatggtgc	actgaaatgg	aaaggggtatt	2150
		2160	2170	2180	2190	2200	
Dnmt3a2	2151	tggcttcccc	gtccactaca	cagacgtctc	caacatgagc	cgcttggcga	2200
DNMT3A2	2151	tggtttccca	gtccactata	ctgacgtctc	caacatgagc	cgcttggcga	2200
		2210	2220	2230	2240	2250	
Dnmt3a2	2201	ggcagagact	gctggggccga	tcgtggagcg	tgccgggtcat	ccgccacctc	2250
DNMT3A2	2201	ggcagagact	gctggggccgg	tcatggagcg	tgccagtcac	ccgccacctc	2250
		2260	2270	2280	2290	2300	
Dnmt3a2	2251	ttcgctccgc	tgaaggaata	ttttgcttgt	gtgtaaggga	catggggggca	2300
DNMT3A2	2251	ttcgctccgc	tgaaggagta	ttttgctgtg	gtgtaaggga	catggggggca	2300
		2310	2320	2330	2340	2350	
Dnmt3a2	2301	aactgaagta	gtgatgataa	aaaagttaaa	caaacaaaca	aacaaaaaac	2350
DNMT3A2	2301	aactgaggta	gcgac-----a	caaagttaaa	caaacaaac-	----aaaaaac	2350
		2360	2370	2380			
Dnmt3a2	2351	aaaacaaaaac	aataaaaacac	caagaacgag			
DNMT3A2	2351	acaaaacat-	aataaaaacac	caagaacatg			

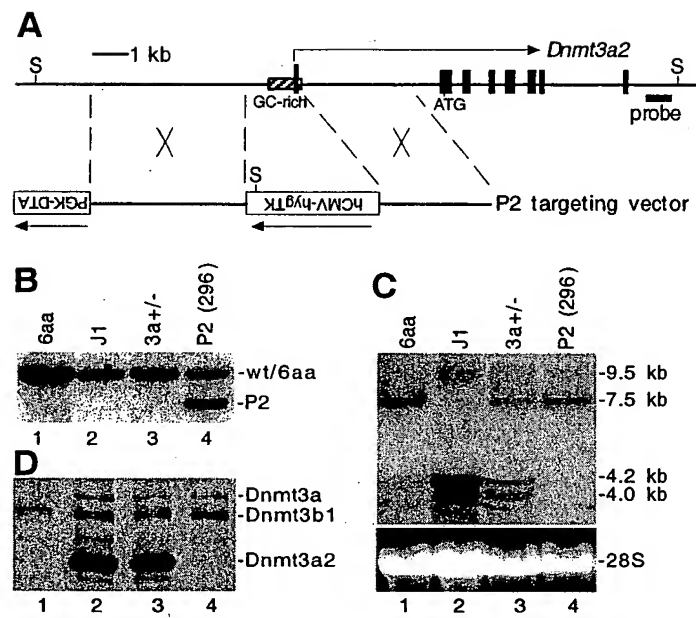
**FIG. 13E-4**

Dnmt3a2	1	MNAVEENQASGESQKVEEASPPAVQOPTDPASPTVATTPEPVGGDAGDKN	50
DNMT3A2	1	MNAVEENQGPGESQKVEEASPPAVQOPTDPASPTVATTPEPVGSDAGDKN	50
Dnmt3a2	51	ATKAADDEPEYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAE	100
DNMT3A2	51	ATKAGDDEPEYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAE	100
Dnmt3a2	101	GTRWVMWFGDGKFSVVCVEKLMPLSSFCSAFHQATYNKQPMYRKATIEVL	150
DNMT3A2	101	GTRWVMWFGDGKFSVVCVEKLMPLSSFCSAFHQATYNKQPMYRKATIEVL	150
Dnmt3a2	151	QVASSRAGKLFPAChDSDESdSGKAVEVQNKQMIEWALGGFOPSGPKGLE	200
DNMT3A2	151	QVASSRAGKLFpVChDSDESdTAKAVEVQNKPMIEWALGGFOPSGPKGLE	200
Dnmt3a2	201	PEEEENPYKEVYTDMMWVEPEAAAYAPPPAKKPRKSTTEKPKVKEIIDE	250
DNMT3A2	201	PEEEENPYKEVYTDMMWVEPEAAAYAPPPAKKPRKSTAEKPKVKEIIDE	250
Dnmt3a2	251	RTRERLVYEVROKCRNIEDICISCGSLNVTLEHPLFIggMCQNCNCFLE	300
DNMT3A2	251	RTRERLVYEVROKCRNIEDICISCGSLNVTLEHPLFVggMCQNCNCFLE	300
Dnmt3a2	301	CAYQYDDDGYQSYCTICCGGREVLmCGNNNCCRCFCVECDLLVGPAAQ	350
DNMT3A2	301	CAYQYDDDGYQSYCTICCGGREVLmCGNNNCCRCFCVECDLLVGPAAQ	350
Dnmt3a2	351	AAIKEDPWNCYmCGHKGTyGLLRREDWPSRLQmFFANNHDOEFDPKvY	400
DNMT3A2	351	AAIKEDPWNCYmCGHKGTyGLLRREDWPSRLQmFFANNHDOEFDPKvY	400
Dnmt3a2	401	PPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVG	450
DNMT3A2	401	PPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVG	450
Dnmt3a2	451	MVRHQGKIMYVGdVRSVtQKHlQEWGPFDLVIGGSPcNDLSiVNPARKGL	500
DNMT3A2	451	MVRHQGKIMYVGdVRSVtQKHlQEWGPFDLVIGGSPcNDLSiVNPARKGL	500
Dnmt3a2	501	YEGTGRLFFEFYRLlHDARPKEGDRPFFWLFENVVAMGVSDKRDISRFL	550
DNMT3A2	501	YEGTGRLFFEFYRLlHDARPKEGDRPFFWLFENVVAMGVSDKRDISRFL	550
Dnmt3a2	551	ESNPVMIDAKEVSAAhRARYfWGNLPGMNRPLASTVNDKLElQECLEHGR	600
DNMT3A2	551	ESNPVMIDAKEVSAAhRARYfWGNLPGMNRPLASTVNDKLElQECLEHGR	600
Dnmt3a2	601	IAKFSKVRTITTRSNSIKQgKDQHfPvFMNEKEDI LWCTEMERVFGFPVH	650
DNMT3A2	601	IAKFSKVRTITTRSNSIKQgKDQHfPvFMNEKEDI LWCTEMERVFGFPVH	650
Dnmt3a2	651	YTDVSNMSRLARQRLlGRSWSVPVIRHLfAPLKEYFACV	689
DNMT3A2	651	YTDVSNMSRLARQRLlGRSWSVPVIRHLfAPLKEYFACV	689

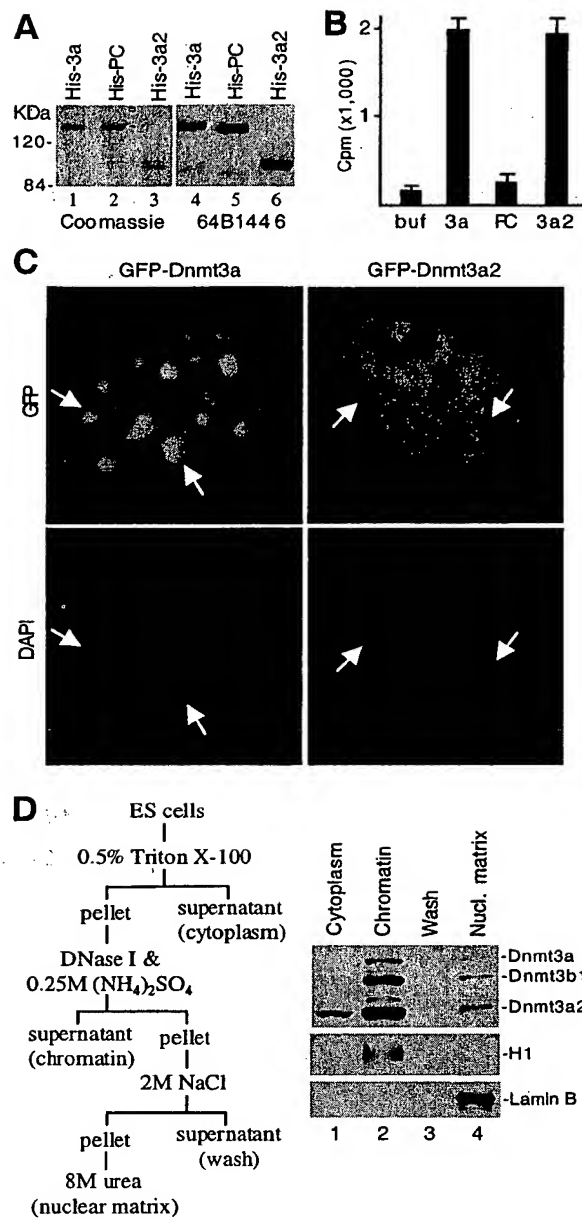
**FIG. 13F**



**FIG. 14**

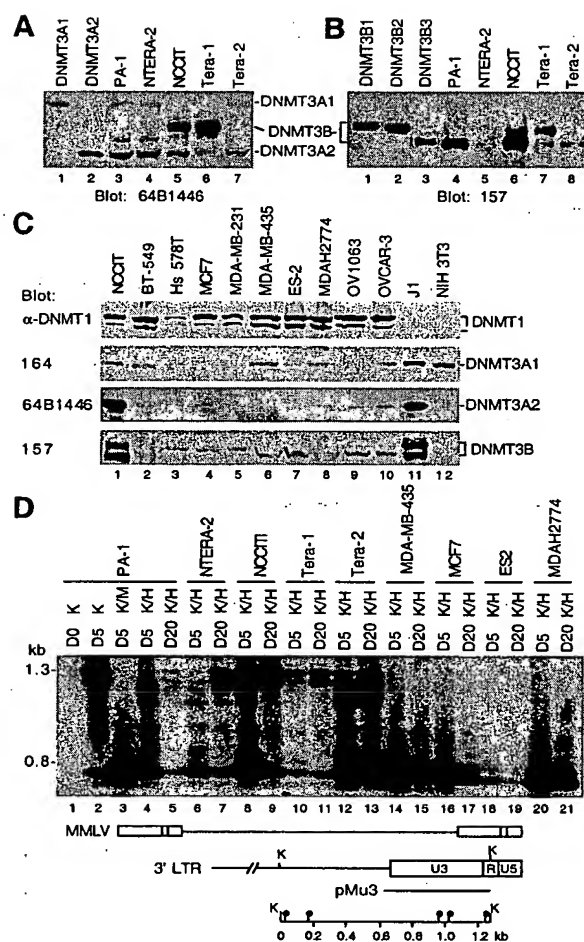


**FIG. 15**

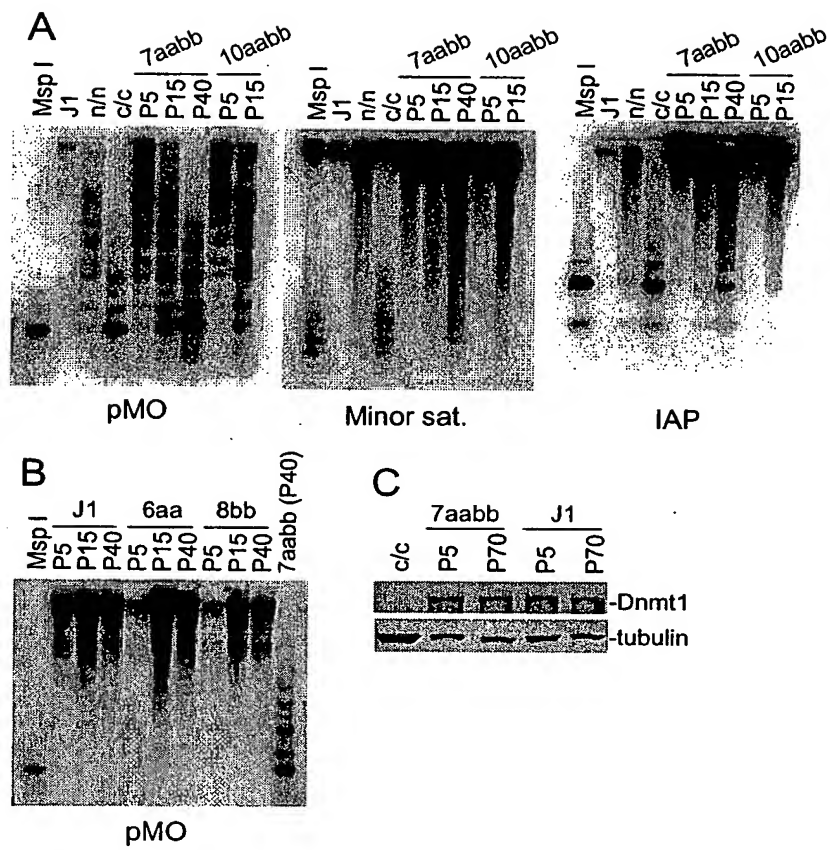


**FIG. 16**

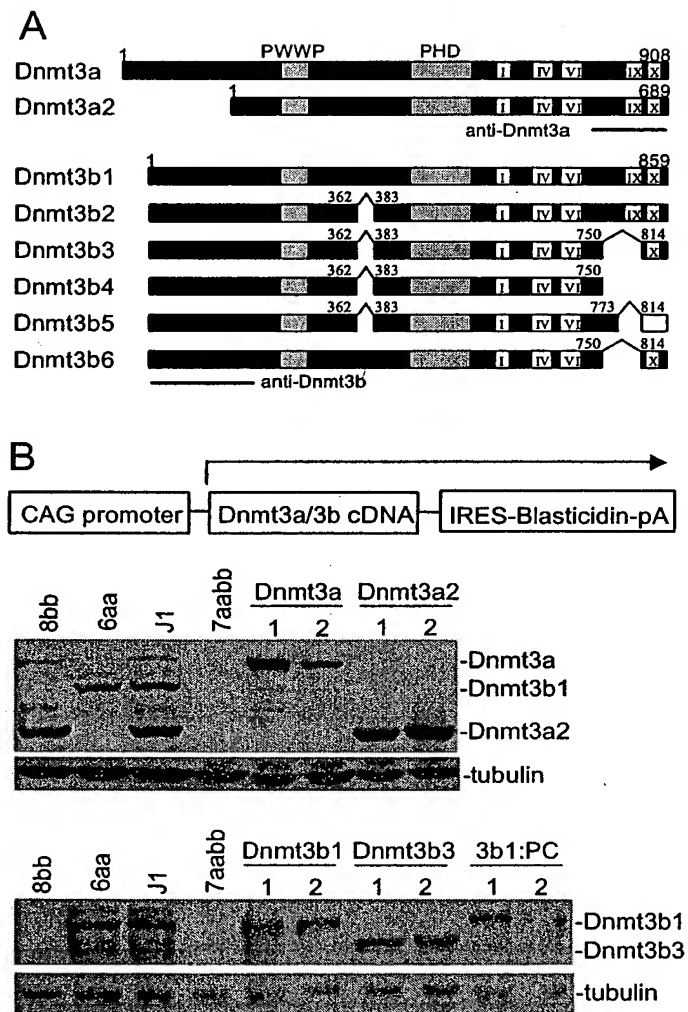




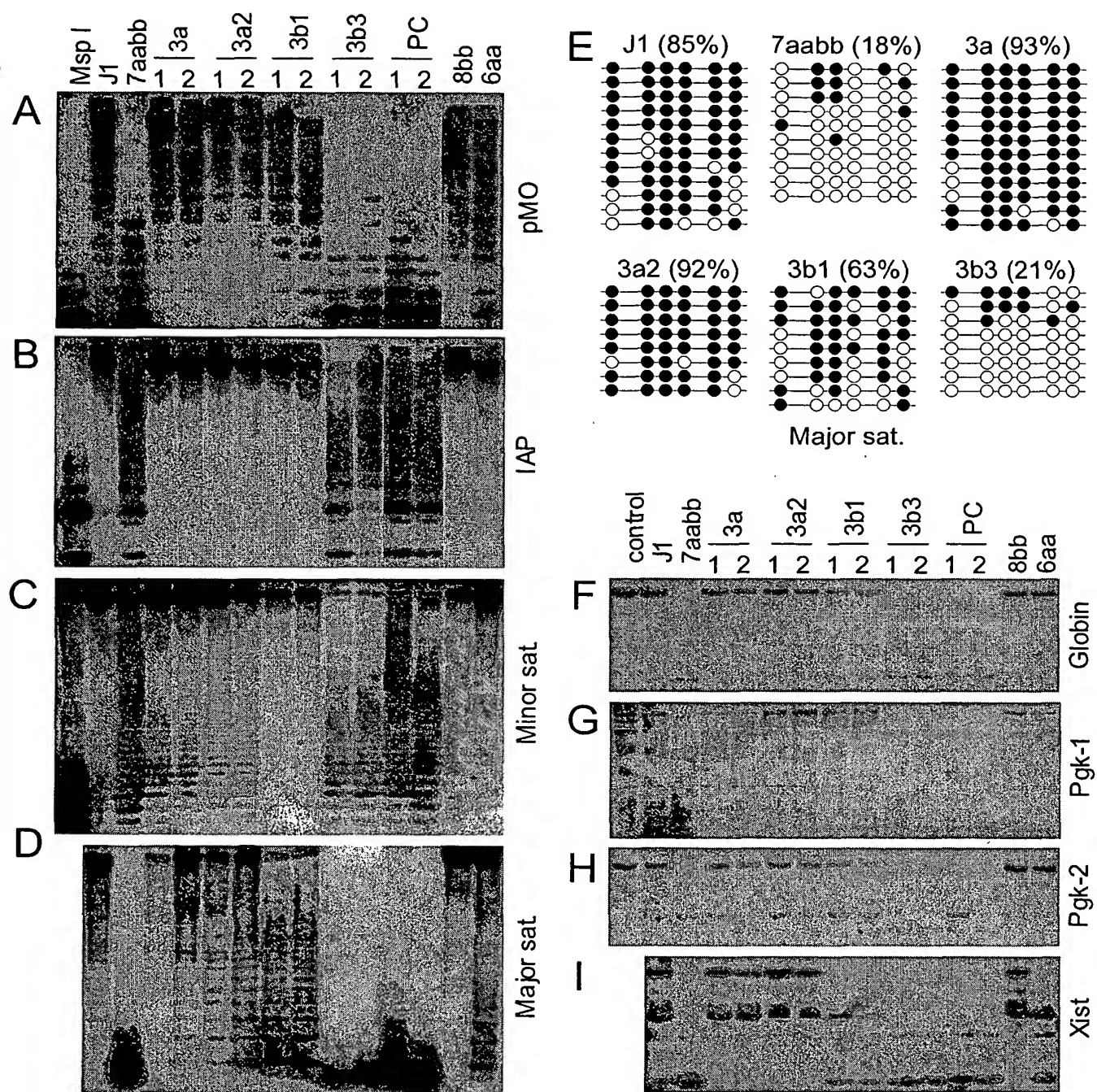
**FIG. 18**



**FIG. 19**



**FIG. 20**



**FIG. 21**

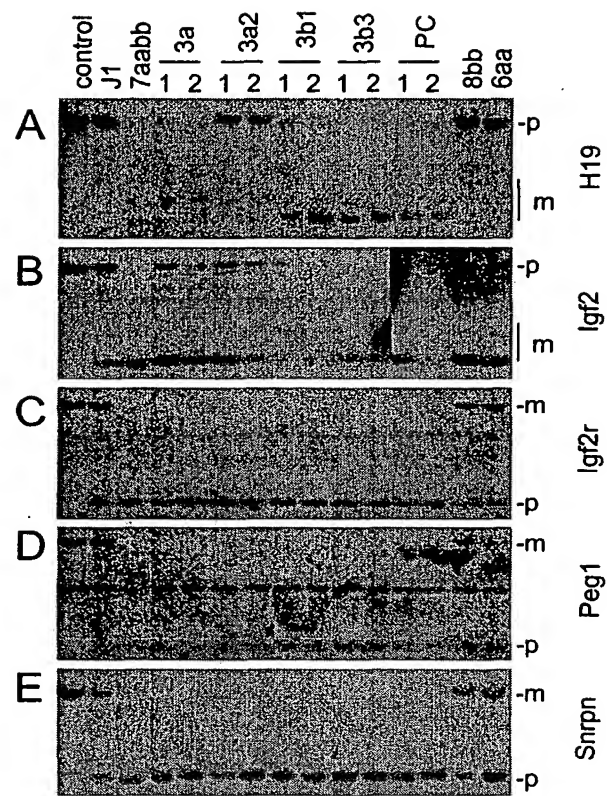
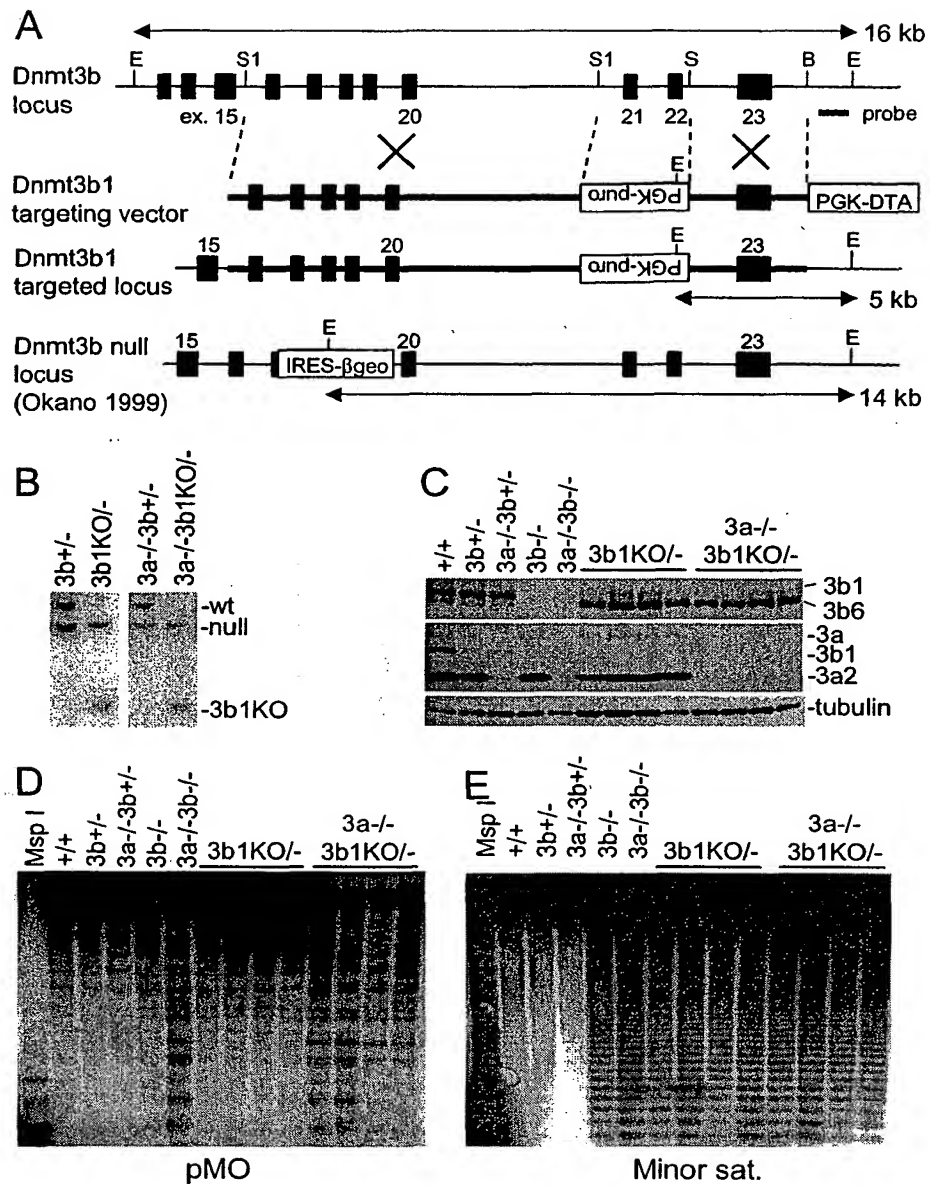
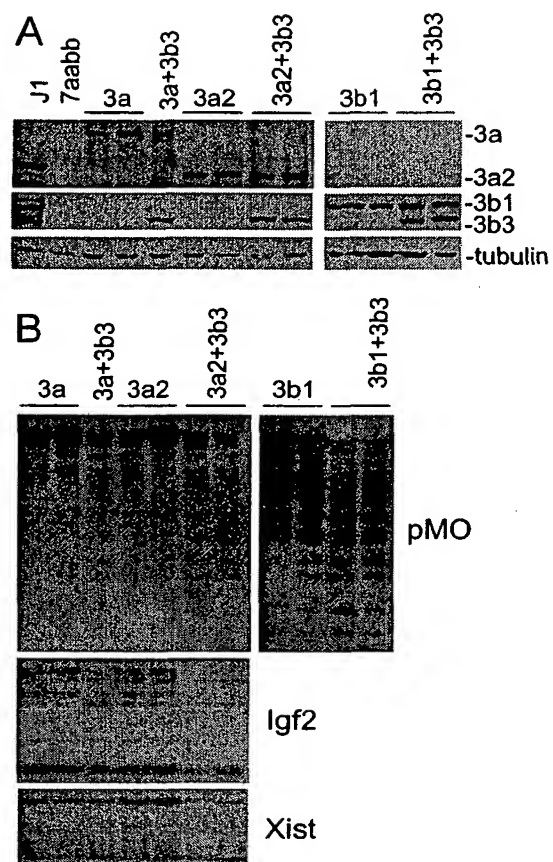


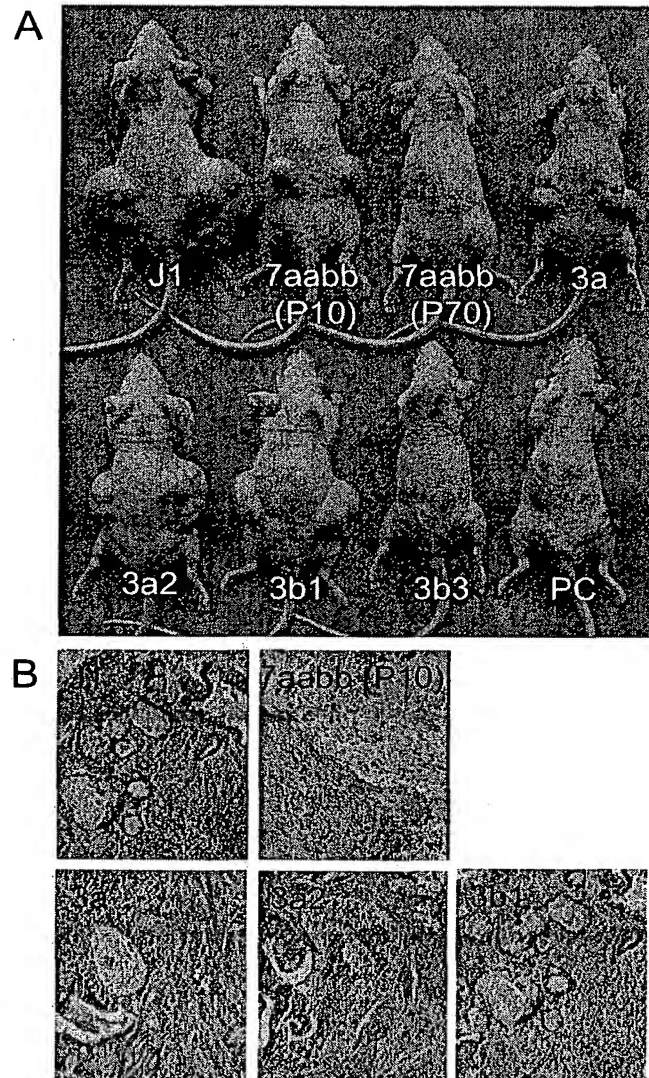
FIG. 22



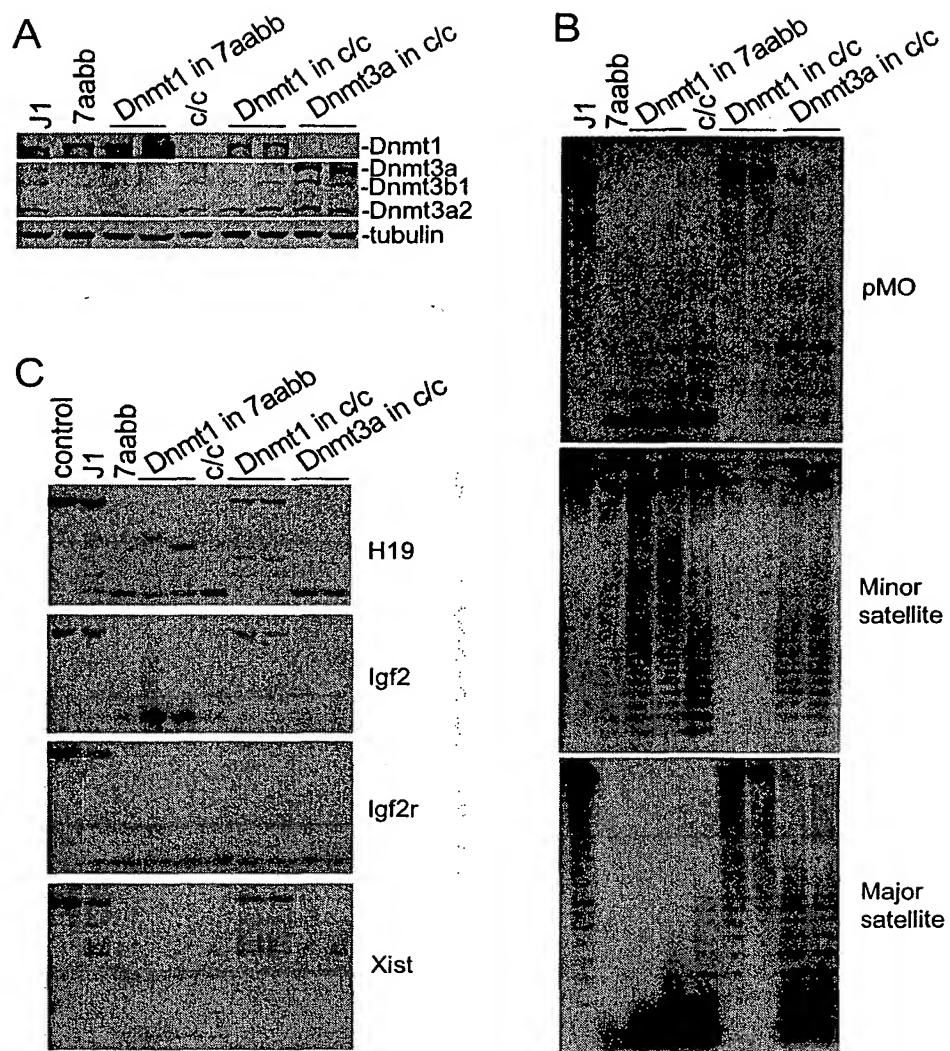
**FIG. 23**



**FIG. 24**



**FIG. 25**



**FIG. 26**

## Mouse Dnmt3a2 Promoter Sequence

```

1  GGAGCCAGGCACCTGGGGTGTTACCTCAGTGCCTTTAGGATATTGGTTTTTCCTAGCTCT
60 AGAGGGCTGATGTCATCACCCCTATTTTGCAGATGAGAAAACAGACATCTTGGGGTTAA
119 GTGGTCTGTGTCAAGGTCACCGCAATGGGATCAGGTCTTCCCCAAGCGTTCAGCCAGA
178 TAGCGGCGGCTCCCTGCTGGGGCATTCTCCTTCAGTTCTTTGTTCTAATTCATCTTGCA
237 AACTTAATCCTGGCTAATCTTTGTAAAATACTCATTACCTTGTTTTTCCAGAACATCT
296 GCCATGTTACAGAATATCTCCATTTCAGTGCTTGACCCCAGTCCCCTACTCAGCCATTT
355 AGCTTTAGTCAAAATTGAGAGGGTGGGTGGAAGAGTTCTTTCTTCCCTTCTACCTGCTTG
414 CCACCTCCAAATCGTGGTTATCTTCTGATCTCTACTGTCTCTATCTCTCACCCACACCCCT
473 TCATTTGATGCAGCCTTCTGCTATCTGCTTGCTGGTTTGGGTAGTTATCCACACAGGAG
532 TTTGCTTTTCAGTGATTCCCCCTTCCCCCACCCCATCTCCCAAGTCTAGTGGAATCTA
591 TCAACTTCCTGAGAGCAGGACCAAGTGTCCATTTCTGTATCCGATGATGCTCCAGTCCT
650 CTAATGGGGGGGGGGGGCGGGGCGCCAGGAGTGGCGTGTGTGCTTCTTCAAACCCAACTT
709 TAGTCCTCTACTGTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
768 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
827 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
886 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
945 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
1004 GAATGGCATGAAATGGCCTACACTTTACCTGGTGGTTCTAGGAGAGAGACACTAGCACG
1063 TGCCTGGGAGTGTGTCTATTACTATTACATAATTGCTGAGACAGGGTTTCGTGATGTTT
1122 AGGCTGGCCTTGAACCTTGTGTTAGTCAAGAATGATCTTAAATTTCTGATCTCTGGTTT
1181 CCAAGTTCTAGGATTACAGGTGTACTTCACCACCAAAGTTTGAACAGCTGCAGATGCC
1240 TTGGCATTGCTCTTAACGAACAGAAAATGAAACAAGCAAGCAAGACCCATTGTGACCCG
1299 GGGGACTCGGGGACTGGACGGGGAAGTTTTCAAAGTCTACTTGTGAACCACGCTTTT
1358 AAGCACCCCTCCATTACCTGTAGCGTGGCGGTGAAGTTATTGTCCTGGGGCGCCCTC
1417 AACCTGCGTGGGACACCTCCTATCCACTCACATCTGTCTTCTGACTTTGCCTAAACTAC
1476 GTTTCGGTAAACTCCGAGCCTCATCTCTAATCTGTAAACTTGCTAGCGCGCTCTCGCAC
1535 GCGCTCTTTTTTTTTTTTTTTTTTCCCGGAAACTCACTTTCTACAACCTTCTCCCCGGAC
1594 TCTCAGGCTGTCTGAAGCCAGCGCTCCTGTCCCACCACCGCTGCTCTGGGTGCCCGCG
1653 GCCCGCACGCACCCTGCCTCCCTCAAGGTCCCCAACTTCCCTATGTACCCCCCATCCC
1712 CAGAGTTGGGGGAAGGGAGCAGAGCGGGCTGTCCCATAAACCTGGCTGGAGGGGCGGGG
1771 CCCTGGGAACGGACTGGCCAGCCTCTCCCCCAGGCCCCCGCGCCCCCTCGGGCCCGGGT
1830 GAGGGGCTGGCCCAGCGCCAGCGTAGGAGGCGGCCCCCTCCCCCGGCCCGCGCTTAG
1889 CCAACCAGAACTCCAGTGGGGCCACGTGACCTGGAGTTCTAGACAAAGAAAATGTTT
1948 CCTCCCTCCCCCGGCGCCCCCTCCCCCTCCCTCTGGCCCCCTCCGCCCCCAACCCCA
2007 ACGCCCCCTGCCCCCTCCCCCAGACGGGCAGCTATTTACAGAGCTTCGGGCCGGGGCTC
2066 ACACCTGAGCTGTACTGCAGAGGGGCTGCACCTGGCCTTATGG

```

FIG. 27

# Human DNMT3A2 promoter sequence

1 GGAGCCAGGCACCTAGAGAATTGTCTCATTGTCATTAGGAGATGGTGGCGTTCCATG  
 60 GCCAAAGAGGGCTGATGTCATCACTCGTTTTGTCAGATGAGACAACAGATTTCTTG  
 119 GGTAAAGTGACTTGTTTAAGGTCATGGTGGTGGAAACAGAACTGAAGTCCAGATCTT  
 178 TTTTTTTTTTTTTTTTTTGGAGACGGAGTCTCGCTCTGTTGCCAGGCTGGAGTGCAGTG  
 237 GCATGATCTCGGCTCACTGCAACATCCGCCTCCTAAGTTGAAGCGATTCTCTTGCCT  
 296 CAGCCTCCCAAGTAGCTGGGATTACTGGCGCACGCCACCACGCCTGGCTAATTTTG  
 355 TATTTTGTAGTAGAGACAAGGTTTCACCATGTTAGTCAGGCCGGTCTCAAACCTCTGA  
 414 CCTCATGATCCGCCTGCCTCAGCCTTCCAAAGTGCTGGGATTATAGGCGTGAGCCAC  
 473 CGCGCTCGGCCAAGTCCAGATCTTCTAACAAGTGCCGCTGCCCAAATAGCCCTCTGC  
 532 TGTGGGGTGCAATTTTCCCTCCATTTCCCTCAGTTCTTCCCTTCTAATTCATCTTGCCAA  
 591 GGCAACTAGGCTGATTTTTCCAAAATACTCATTTCATCTTGTGAGAAAACCTGCGGTT  
 650 ATTCTTCCCTGCTACAGAATATACCCAAGGACGCACCTGAAGGCTTGCCATTACCTT  
 709 GCCCTGTCGTGTAAGTGGGAGGGTGGAGGTGGGCGAGGGTCTCCTCCCTCCCCAGCCC  
 768 GGCAGCTCTTGCTCATCCTACCCATCTCACCTCATTCCAAGTCCGATCCAGCCTCCA  
 827 GGCCAGTCGGCTCACCTGGAAGTACCTCTGACCTCTTTTGTATCCATGCCGCC  
 886 ATTTTTTTTCTACTTGGTATTTGTGGCATAGTTACCTTTACATATGTTTGTTTTACAG  
 945 TGATCCTTTCATATTTCTCCAAGTCTAGTGAATCTTCAACCCCTCGAGGGCAGAGC  
 1004 CAACAGGGTCTATTTCTTTATCTGATCCTACAGCCAACGTAATGGAGGGCTGTGGGT  
 1063 GGGGACTGCGTCTGCCTTGGGGGTAGGTGCCTTTGTTTCAGGAGGAGGAAGCTTGAAA  
 1122 TGGCGGAGGCTGCACCTGGAGGCCGCACCTGGAGGCCCCAGGAGAGGAGTCAGGTCT  
 1181 TCTCGATCTGCAGATGTTTGAGCCTGGGAATGAAGGAATTGCTGAACCTTCTGAAGG  
 1240 AGCGCCCTCGCCGCGACCAACCTTGCAAACAGGAAAATGAGAAATCCAGGGAAGGCC  
 1299 CAGAGTGACGCAGGGGCCCTGGGACTCGAAGCCTGACCTCCTCACGCCGCGCTTTTT  
 1358 GAGGCCCCCCCCGCTTCTCTATTACCTGTAGTGTGGAGGCGGGAGACCCCCCAAACA  
 1417 ATCCCCGATCTGGAGCGCTCCCAATGCCTGCGCGCGCTGCTGTCACTCTCCGTCTG  
 1476 TGTGCTGAGTTTTCTACAGCTTCTTGGGCCTCCTATCTGTAAGCTTTTTCTTTTTT  
 1535 TTTTTTTGGTTGTGCTTTCAGAGAACTCACTTTTCACAACTTCTCCCGGCTCTCCC  
 1594 AGGCCGTCCGAAAGCTCCGGCTTGCTTTGCCCCGACCCCCCGGCTCCCTCCGGGCAG  
 1653 GCGGCTCGGGAGCAGCCCCCTTCCCTCCCCCTCCCGGCCCGCCCGGCTAATCT  
 1712 CTTCCAGAGCTGGGGGAGGGGCCAGGCGGTCTTCCCGAAGGCGGGGCGCTCCCTGCA  
 1771 GCGCGGCTGGGCGGGCCCTGGGAACGGGCGGGGAACGGCCTCGCCCCCGGCCCCG  
 1830 CGCCCCCTCGGACCGGAGAAGAGGGGCTGGCCAGCGCCAGCGTCGGAGCGCCGGCCC  
 1889 CCTCCCCGGGCGGCTCGCAGCCAACCAGGCCCTCCAGCGGGGCCCACGTGACCTGGA  
 1948 GTCCTAGACAAAGAAAATGTTCCCTCCCTCCCCCGCGCCCCCTCCCTCCAG  
 2007 TGGCCCCCTCCGCCCCCAGCCCCATCGCCCCCTTCCCTCCCCCAAGACGGGCAGCT  
 2066 ACTTCCAGAGCTTCAGGGCCGCGGCTCACACCTGAGCGCGACTGCAGAGGGGCTGCA  
 2125 CCTGGCCTTATGG

FIG. 28

# Mouse and human Dnmt3a2 promoter alignment

Top Sequence = mouse Dnmt3a2 promoter, 1858 bp (gap not counted)  
Bottom Sequence = human DNMT3A2 promoter, 2065 bp

1-104 (1-105) 77% ==  
1289-1338 (1475-1530) 82% ==  
1518-1858 (1724-2065) 87%

```

0      :      :      :      :      :      :      :      :
1  GGAGCCAGGCACCTGGGGTGTACCTCAGTGCCTTTAGGATAT  TGGTT
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
1  GGAGCCAGGCACCTAGAGAATTGTCTCATTGTATTAGGAGATGGTGGCG

50
49  TTCC TAGCTCTAGAGGGCTGATGTCATCACCCCTATTTGCAGATGAGA
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
51  TTCCATGGCCAAAGAGGGCTGATGTCATCACTCGT  TTTGCAGATGAGA

100
98  AAACAGA
   |||||
99  CAACAGA

0      :      :      :      :      :      :      :      :
1289 TCTTTTTTTTTTTTTT TT T  TCCCG GAAACTCAC TTTCTACAACT
     ||||| ||||| ||||| ||||| ||||| ||||| |||||
1475 TCTTTTTTTTTTTTTT TGGTGTGCTTCAGAGAACTCACTTTTC ACAACT

50
1332 TTCTCCC
     |||||
1524 TTCTCCC

0      :      :      :      :      :      :      :      :
1518 GGGCCCTGGGAACGGAC TGG CCAGCCTCTCCCCCAGGCCCGCGCGCC
     ||||| ||||| ||||| ||||| ||||| ||||| |||||
1724 GGGCCCTGGGAACGGGCGGGGAACGGGCTCGCCCCCGGCCCG  GCGCC

50
1566 CCTCGGGCCCG GGTGAGGGGCTGGCCCGAGCGCGTAGGAG GCCGG
     ||||| ||||| ||||| ||||| ||||| ||||| |||||
1772 CCTCGGACCGGAGAAGAGGGGCTGGCCCGAGCGCGTAGGAGCGGCGCGCG

100
1614 CCCCCTCCCCCGGCC CGCGCTTAGCCAACCAGAACTCCAGTGGGGCC
     ||||| ||||| ||||| ||||| ||||| ||||| |||||
1822 CCCCCTCCCCCG GGCCGCTCGC' AGCCAACCAGGCCCTCCAGCGGGGCC

150
1663 CACGTGACCTGGAGTTCTAGACAAAGAAAATGTTCCCTCCCTCCCCCGG
     ||||| ||||| ||||| ||||| ||||| ||||| |||||
1869 CACGTGACCTGGAGTCTTAGACAAAGAAAATGTTCCCTCCCTCCCCCGG

200
1713 GCGCCCCC TCCCCTCCCTCTGGCCCCCTCCGCCCCCAACCCCAACGCC
     ||||| ||||| ||||| ||||| ||||| ||||| |||||
1919 CCGCCCCCTCCCC TCCCAGTGGCCCCCTCCGCCCCCAGCCCCATCGCC

250
1762 CCCTGCCCCCTCCCCCAGACGGGCAGCTATTTACAGAGCTTC GGGCCGG
     ||||| ||||| ||||| ||||| ||||| ||||| |||||
1968 CCCTTCCCCCTCCCCCAAGACGGGCAGCTACTTCCAGAGCTTCAGGGCCGC

300
1811 GGCTCACACCTGAGC TGTACTGCAGAGGGGCTGCACCTGGCCTTATGG
     ||||| ||||| ||||| ||||| ||||| ||||| |||||
2018 GGCTCACACCTGAGCGCG ACTGCAGAGGGGCTGCACCTGGCCTTATGG

```

FIG. 29